

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2005, 12:04:33 ; Search time 39 Seconds  
(without alignments)  
463.814 Million cell updates/sec

Title: US-10-777-053-40  
Perfect score: 1012  
Sequence: 1 MNGDDAFARRPTVGAQIPEK.....RRKQLVIYEISDPEEDDE 188

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_79:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1012	100.0	188	2 S55058	synovial sarcoma X hum
2	772	76.3	188	2 S55057	SSX1 protein - hum
3	107	10.6	1111	2 T00324	hypothetical prote
4	101	10.0	289	2 A48913	KRAB zinc finger p
5	99	9.8	392	2 T19327	hypothetical prote
6	94.5	9.3	952	2 S64473	translation initia
7	91.5	9.0	252	2 T20321	hypothetical prote
8	91.5	9.0	1507	2 B47328	natural killer cel
9	90.5	8.9	1153	2 T21386	hypothetical prote
10	88.5	8.7	213	2 JQ0075	neuromodulin - gol
11	88.5	8.7	270	2 B89734	protein F32E10.2 [
12	88	8.7	445	2 T23908	hypothetical prote
13	87.5	8.6	186	2 S30221	nonhistone chromos
14	87.5	8.6	209	1 NSH02	nonhistone chromos
15	87.5	8.6	541	2 T48836	hypothetical prote
16	87.5	8.6	586	2 B84434	hypothetical prote
17	87	8.6	346	2 C88961	protein F59A7.5 li
18	87	8.6	1559	2 T30535	calcium channel al
19	86.5	8.5	654	2 A57785	finger protein ZNF
20	86	8.5	1087	1 QFMSH	neurofilament trip
21	85.5	8.4	990	2 I51618	nucleolar phosphop
22	85.5	8.4	2251	2 B54972	voltage-dependent
23	85.5	8.4	2270	2 A54972	voltage-dependent
24	85	8.4	210	2 A34719	nonhistone chromos
25	85	8.4	210	2 S54774	high mobility grou
26	85	8.4	1162	2 T51040	hypothetical prote
27	85	8.4	1167	1 A35066	streptococcal C5a
28	84.5	8.3	755	2 S32103	filensin - bovine
29	84	8.3	1356	2 T16754	hypothetical prote

30	84	8.3	1647	2 S45252	SNF2beta protein -
31	83.5	8.3	301	2 E29149	proline-rich prote
32	83.5	8.3	595	2 JC7779	Kruppel-associate
33	83.5	8.3	636	2 I48689	gene NK10 protein
34	83.5	8.3	920	2 T52426	dynammin-like prote
35	83.5	8.3	1280	2 T00365	hypothetical prote
36	83.5	8.3	1626	2 A39242	DNA topoisomerase
37	83.5	8.3	1663	2 T42092	s-afadin - rat
38	83.5	8.3	1829	2 T41751	l-afadin - rat
39	83	8.2	783	2 F84514	hypothetical prote
40	83	8.2	1062	2 T14151	inv protein - mous
41	83	8.2	1110	2 I51116	NF-180 - sea lamp
42	83	8.2	1235	2 T17457	SARA protein - Afr
43	83	8.2	1948	2 S00485	gene 11-1 protein
44	82.5	8.2	1069	2 S27922	nuclear antigen EB
45	82	8.1	279	2 T15662	hypothetical prote

ALIGNMENTS

RESULT 1

SS5058  
synovial sarcoma X chromosome breakpoint protein SSX2 - human  
C:Species: Homo sapiens (man)  
C>Date: 28-Oct-1996 #sequence revision 13-Mar-1997 #text\_change 09-Jul-2004  
C:Accession: S55058; I68675; S46270  
R:Crew, A.J.; Clark, J.; Fisher, C.; Gill, S.; Grimer, R.; Chand, A.; Shipley, J.; Guste  
EMBO J. 14, 2333-2340, 1995  
A:Title: Fusion of SYT to two genes, SSX1 and SSX2, encoding proteins with homology to t  
A:Reference number: S55057; MUID:95292974; PMID:7539744  
A:Accession: S55058  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-188 <CRE>  
A:Cross-references: UNIPROT:Q16385; EMBL:X86175; NID:G829113; PIDN:CAA60111.1; PID:G42184  
R:de Leeuw, B.; Bailemans, M.; Oide Weghuis, D.; Geurts van Kessel, A.  
Hum. Mol. Genet. 4, 1097-1099, 1995  
A:Title: Identification of two alternative fusion genes, SYT-SSX1 and SYT-SSX2, in t(X;11  
A:Reference number: I54381; MUID:95384157; PMID:7655467  
A:Accession: I68675  
A>Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 111-188 <RES>  
A:Cross-references: GB:S79332; NID:G1087049; PIDN:AAB35379.1; PID:G1087050  
R:Clark, J.; Rocques, P.J.; Crew, A.J.; Gill, S.; Shipley, J.; Chan, A.M.L.; Gusterson, F  
Nature Genet. 7, 502-508, 1994  
A:Title: Identification of novel genes, SYT and SSX, involved in the t(X;18) (p11.2;q11.2  
A:Reference number: S46269; MUID:95038836; PMID:7951320  
A:Accession: S46270  
A:Molecule type: mRNA  
A:Residues: 111-188 <CLA>  
A:Cross-references: EMBL:X79200  
C:Genetics:  
A:Gene: SYT-SSX2

Query Match 100.0%; Score 1012; DB 2; Length 188;  
Best Local Similarity 100.0%; Pred. No. 2.4e-78;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MNGDDAFARRPTVGAQIPEKIQAFDDIAKYESKEWEKKKASEKIFYYVMKKYEAMTK	60
Db	1	MNGDDAFARRPTVGAQIPEKIQAFDDIAKYESKEWEKKKASEKIFYYVMKKYEAMTK	60
Qy	61	LGFKATLPPFCMCKAEAFQGNLDNDPNRGVQVERPQMTFGRLQGISPKIMPKKPAEG	120
Db	61	LGFKATLPPFCMCKAEAFQGNLDNDPNRGVQVERPQMTFGRLQGISPKIMPKKPAEG	120
Qy	121	NDSEVPASGQNDGKELCPGKPTTSEKIHRSQPKRGEHAWTHRLRERKQLVIYEI	180
Db	121	NDSEVPASGQNDGKELCPGKPTTSEKIHRSQPKRGEHAWTHRLRERKQLVIYEI	180
Qy	181	SDPEEDDE	188

Q44431 N363

Db 181 SDPEEDDE 188  
|||||  
RESULT 2  
S55057  
C:Species: Homo sapiens (man)  
C:Date: 23-Aug-1995 #sequence\_revision 19-Oct-1995 #text\_change 09-Jul-2004  
C:Accession: S55057; 154381  
R:Crow, A.J.; Clark, J.; Fisher, C.; Gill, S.; Grimer, R.; Chand, A.; Shipley, J.; Guste  
EMBO J. 14, 2333-2340, 1995  
A:Title: Fusion of SYT to two genes, SSX1 and SSX2, encoding proteins with homology to t  
A:Reference number: S55057; MUID:95292974; PMID:7539744  
A:Accession: S55057  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-188 <CRE>  
A:Cross-references: UNIPROT:Q16384; EMBL:X86174; NID:g829112; PIDN:CAA60110.1; PID:g4218  
R:de Leeuw, B.; Bailemans, M.; Olde Weghuis, D.; Geurts van Kessel, A.  
Hum. Mol. Genet. 4, 1097-1099, 1995  
A:Title: Identification of two alternative fusion genes, SYT-SSX1 and SYT-SSX2, in t(X;1  
A:Reference number: 154381; MUID:93384157; PMID:7655467  
A:Accession: 154381  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 111-188 <RES>  
A:Cross-references: GB:S79325; NID:g1087047; PIDN:AAB35378.1; PID:g1087048  
C:Genetics:  
A:Gene: SYT-SSX1  
Query Match 76.3%; Score 772; DB 2; Length 188;  
Best Local Similarity 78.2%; Pred. No. 4.4e-58;  
Matches 147; Conservative 12; Mismatches 29; Indels 0; Gaps 0;  
Qy 1 MNGDDAFRRPTVGAQIPEKIQAFDDIAKYSKEWEKMKASEKIFVYMKRYEAMTK 60  
|||||  
Db 1 MNGDDTFAKRPDDAKASEKRSKAPDDIATYFSKKEWKMKYSEKISVYMKRYNKAATK 60  
Qy 61 LGFKATLPFFCNKAEDFQGNLDNDNRGNQVERPQMTFGRLOGISPKIMPKPAEEG 120  
Db 61 LGFKVTLTFFCNKQATDFQGNDFNDNRRIQVEHPQMTFGRHLRIIPKIMPKPAEDE 120  
Qy 121 NDSEVPASGPQNDGKELCPGKPTTSEKIHRSKPRGHEAWTHRLRERKQLVYEEI 180  
Db 121 NDSKGVSEASGPQNDGKQLHPPGKANISEKINKRSKPRGKHAWTHRLRERKQLVYEEI 180  
Qy 181 SDPEEDDE 188  
Db 181 SDPEEDDE 188  
RESULT 3  
T00324  
hypoetical protein KIAA0543 - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 09-Jul-2004  
C:Accession: T00324  
R:Nagase, T.; Ishikawa, K.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Ohara, O.  
DNA Res. 5, 31-39, 1998  
A:Title: Prediction of the coding sequences of unidentified human genes. IX. The complet  
A:Reference number: Z14086; MUID:98290545; PMID:9628581  
A:Accession: T00324  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1111 <NAG>  
A:Cross-references: UNIPROT:O60290; EMBL:AB011115; NID:g3043609; PIDN:BAA25469.1; PID:g3  
A:Experimental source: brain  
C:Genetics:  
A:Note: KIAA0543  
Query Match 10.6%; Score 107; DB 2; Length 1111;  
Best Local Similarity 24.2%; Pred. No. 0.51;

Matches 44; Conservative 34; Mismatches 70; Indels 34; Gaps 9;  
Qy 16 QIP---EKIQAFDDIAKYSKEWEKMKASEKIFVYMKRYEAMTKLGFKATLPFFM 71  
Db 265 EVPVVFELPVFEDVAVYFTREWGMLDKKQKELRDVMMNYELLASLGPAAKPDLI 324  
Qy 72 --CNKRAEDFQGNLDNDNRGNQVERPQMTFGRLOGISPKIMPKPAEAGNDSSEVPEA 129  
Db 325 SKLERRAAPW-----IKDPN-----GPKWKGGRPPG-NKQWAVREADTQSAADSALL 372  
Qy 130 SPQNDGKELCPGKPTTSEKIHRSKPRGHEAWTHRLRERK-----QIVVYEEISDP 183  
Db 373 PGSPVEARASC-----CSSSICERGDGPRRIKRTYRPRSIORSWFGQPFWLVI-----DP 422  
Qy 184 EE 185  
Db 423 KE 424  
RESULT 4  
A48913  
KRAB zinc finger protein 75 - human  
N:Alternate names: ZNF75 protein  
C:Species: Homo sapiens (man)  
C:Date: 10-May-1996 #sequence\_revision 10-May-1996 #text\_change 09-Jul-2004  
C:Accession: A48913; A43288; S47339  
R:Villa, A.; Zucchi, I.; Pillia, G.; Strina, D.; Susani, L.; Morali, F.; Patrosso, C.; Fra  
Genomics 18, 223-229, 1993  
A:Title: ZNF75: isolation of a cDNA clone of the KRAB zinc finger gene subfamily mapped i  
A:Reference number: A48913; MUID:94116987; PMID:8288223  
A:Accession: A48913  
A:Status: preliminary; nucleic acid sequence not shown; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-289 <RES>  
A:Cross-references: UNIPROT:P51815; GB:S67970; NID:g460902; PIDN:AAB29696.1; PID:g460903  
A:Experimental source: lung fibroblasts  
R:Villa, A.; Patrosso, C.; Biunno, I.; Frattini, A.; Repetto, M.; Mostardini, M.; Evans,  
Genomics 13, 1231-1236, 1992  
A:Title: Isolation of a zinc finger motif (ZNF75) mapping on chromosome Xq26. .  
A:Reference number: A43288; MUID:92372018; PMID:1505955  
A:Accession: A43288  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 71-289 <VIL>  
A:Cross-references: GB:S43109; NID:g254027; PIDN:AAB22971.1; PID:g254028  
A:Note: sequence extracted from NCBI backbone (NCBIN:111827, NCBI:P:111830)  
R:Marino, M.; Archidiacono, N.; Franze, N.; Rosati, M.; Rocchi, M.; Ballabio, A.; Grimalc  
submitted to the EMBL Data Library, July 1992  
A:Description: The human zinc finger gene family: isolation, mapping and expression of t  
A:Reference number: S47339  
A:Accession: S47339  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 139-289 <MAR>  
A:Cross-references: EMBL:X68010; NID:g525243; PIDN:CAA48147.1; PID:g525244  
C:Genetics:  
A:Gene: GDB:ZNF75  
A:Cross-references: GDB:131726; OMIM:314997  
A:Map position: Xq26-Xq26  
A:Introns: 11/3  
C:Keywords: zinc finger  
Query Match 10.0%; Score 101; DB 2; Length 289;  
Best Local Similarity 22.2%; Pred. No. 0.34;  
Matches 46; Conservative 22; Mismatches 65; Indels 74; Gaps 9;  
Qy 17 IPEKIQ-KAFDDIAKYSKEWEKMKASEKIFVYMKRYEAMTKLGFKATLPFFMCK 74  
Db 7 LPESLSLLTFEDVAVYFSEEEWQLNPLEKTLNDVMDIYETVISLGLKLGK----- 58  
Qy 75 RAEDFQGNLDND-----PNRGNOVERPQMTFGR-----LQGI 107  
Db 59 -----NDTGNADHPISVSTSEIQTSGCEVSKTKRMKTAQKTMGRNPGDTHSVQKWHRA 111

```

QY      108  SPKIMPKPAEAGNDSEEVPE-----ASGP-----QNDGKELCPPGKPTTSEKIH  154
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      112  PFRKKKKKFA-----TCKQELPKMLDLHGSGTGKPFKCEQCGKSFVSSDLNKHRIHTG  168
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      155  SGPGR-----GSHAWTHR  167
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      169  EKPYKCOQCDRRFRWSSDLNKHFMTHQ  195
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 5  
T19327  
hypothetical protein Cl6C10.6 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T19327  
R:Lloyd, C.  
submitted to the EMBL Data Library, November 1994  
A:Reference number: Z19108  
A:Accession: T19327  
A:Status: preliminary; translated from GE/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-392 <MIL>  
A:Cross-references: UNIPROT:Q09252; EMBL:Z46787; PIDN:CAA86744.1; GSPDB:GN00021; CESP:CL16C10.6  
A:Experimental source: clone Cl6C10  
C:Genetics:  
A:Gene: CESP:CL6C10.6  
A:Map position: 3  
A:Introns: 37/3; 87/3

Query Match	9.8%;	Score 99;	DB 2;	Length 392;
Best Local Similarity	23.9%;	Pred. No. 0.72;		
Matches	49;	Conservative	28;	Mismatches 80;
				Indels 48;
				Gaps 8;

  

QY	5	DAFARRPTVGAQIPEKIQAFFD	----	IAKYFSKEWE	----	KMKASEKIFVYMKRKY	55
		: : :		: : :		: :	
DB	116	EQSRERQQLREREKEGDFD	KEVFTVCAYRKQEEVKHREQEAEAAAFNDMTSVOK	175			
		: :		: :		: :	
QY	56	EAMTKLGFKATLPFFMCKRAED	FQGNLDLNDNRGNQVERPQMTFGRLOGISPKMP	-- 113			
		: :		: :		: :	
DB	176	QKLWEIGMGRTLL	-----	NDLARDPTAIKQRKEQKNVRKREDSDEEIDPKP	222		
		: :		: :		: :	
QY	114	----	KPAEB	----	GNDSEEVPEASGPONDGKELCPGKPTTSEKIHERSGPKGEHA	163'	
		: :		: :		: :	
DB	223	EKSDKPAELKKSIVSDSDEKA	PKPPQKNFEGDLKPGLNTVSKK	-----	KATTHA	275	
		: :		: :		: :	
QY	164	WTHRLRERKQLVYEEISDPEEDD	188				
		: :		: :		: :	
DB	276	--	ERIRORNYTP	-----	TPSSSDDE	293	
		: :		: :		: :	

RESULT 6  
S64473  
translation initiation factor eIF-4F TIF4631 - yeast (*Saccharomyces cerevisiae*)  
N;Alternate names: cap-binding protein complex chain TIF4631; protein G7036; protein YGK1  
C;Species: *Saccharomyces cerevisiae*  
C;Date: 17-May-1996 #sequence\_revision 17-May-1996 #text\_change 09-Jul-2004  
C;Accession: S64473; A48086  
R;Rieger, M.; Mueller-Auer, S.; Brueckner, M.; Schaefer, M.  
submitted to the Protein Sequence Database, May 1996  
A;Reference number: S64071  
A;Accession: S64473  
A;Molecule type: DNA  
A;Residues: 1-952 <RIE>  
A;Cross-references: UNIPROT:P99935; EMBL:Z72947; NID:G1323279; PID:e243549; PID:gl323284  
A;Experimental source: strain S288C  
R;Goyer, C.; Altmann, M.; Lee, H.S.; Blanc, A.; Deshmukh, M.; Woolford Jr., J.L.; Trachsel  
Mol. Cell. Biol. 13, 4860-4874, 1993  
A;title: TIF4631 and TIF4632: two yeast genes encoding the high-molecular-weight subunit  
quence and carry out an essential function.  
A;Reference number: A48086; MUID:93330281; PMID:8336723  
A;Accession: A48086  
A;Molecule type: DNA

A:Residues: 1-6, 'Q', 8-36, 'N', 38-109, 'K', 111-206, 'K', 208-360, 'E', 362-952 <G>  
A:Cross-references: GB:IL16923; NID:G295674; PIDN:AAA02757.1; FID:G295675  
C:Genetics:  
A:Gene: SGD: TIF4631  
A:Cross-references: SGD:S0003394; MIPS:YGR162W  
A:Map position: 7R  
C:Keywords: heterodimer; protein biosynthesis

Query Match	9.3%	Score 94.5;	DB 2;	Length 952;
Best Local Similarity	22.5%;	Pred. NO. 4.8;		
Matches	45; Conservative 31;	Mismatches 69;	Indels 55;	Gaps 6
Qy	10 RPTVGAIQPEIKIOKAFDDI-----AKYFSKEEWEKMKASEKIFVYVMKRKYE	56		
Db	159 RSTVSPQESKLKETS DSTSTSTTPTPSTNDSKASSEENISEAEKTRNFIEQVKLRKA	218		
Qy	57 AMTKLGFATLPPECNKNRAEDFOGNDLDNDPRNGNQVERPQMTGRLOGISPKIMPKKP	116		
Db	219 ALEK-----KRKEQLEGSSGN-----IPMKTTPENV	246		
Qy	117 AEENGSEEVPEASGP-----QNDGKELCPGPKTTSEK--IHRS GPKRCEHAWTHRL	168		
Db	247 EEKSDPEVTETKTPAEKSAPVEVKQETPAEEGEQCKQIKEESTPK--VLTFEAERL	304		
Qy	169 RERKOLVIYEISDPDEDE	188		
Db	305 KLIKQOEREKTEKGNKKE	324		

```

RESULT 7
T20321
hypothetical protein D1086.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T20321
R:Smyle, R.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19257
A:Accession: T20321
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-252 <WIL>
A:Cross-references: UNIPROT:O17724; EMBL:Z81491; PIDN:CAB04017.1; GSPDB:GNO
A:Experimental source: clone D1086
C:Genetics:
A:Gene: CESP:D1086.4
A:Map position: 5
A:Introns: 16/1; 49/1; 92/3; 169/1; 200/1; 230/3

```

	Query Match	9.0%; Score 91.5; DB 2; Length 252;
	Best Local Similarity	20.8%; Pred. No. 1.9;
	Matches	45; Conservative 37; Mismatches 57; Indels 77; Gaps 12;
Qy	12 TVGAGQ---IPEK-----IQAFDDIAKYFSKEWEKMKASEKIFVYVMKR---KYEAMTK 60	
Dd	57 SVGARFYSLFQKEQDEVNFAFNSRLRYLKRE-----PYAKLRKANSKSVSKPK 107	
Qy	61 LGFKATLPPFWCNKRAEDFGCNDLDNDPNRGN-----OVERPQMTFGKLQ----- 105	
Dd	108 -----EREKSVDNSDEADRNGKNKKTQKNASKNCQIEKSSNNSGILKKSGS 154	
Qy	106 GISPKIMPKKPA-----EGDNDSBVP-----EASGPQDNGKELC--PPCKP 145	
Dd	155 GISVASPKKSVAPAGVYEDLTDDLEFNLSVINSDRPTSQCDNPARRMCGRPPTYKH 214	
Qy	146 TTSEKIHERGGPKRGE-----HA-----WTHRLRERK 172	
Dd	215 RDTESOEITGSKKOKIFPTPHKEKPAAWWSFRIPKKR 250	

RESULT 8  
B47328  
natural killer cell tumor-recognition protein - mouse

N:Alternate names: cyclophilin-related NK-tumor recognition protein; natural killer-tumor  
C:Species: Mus musculus (house mouse)  
C:Date: 17-Jul-1994 #sequence\_revision 17-Jul-1994 #text\_change 09-Jul-2004  
C:Accession: B47328; 177662  
R:Anderson, S.K.; Gallinger, S.; Roder, J.; Frey, J.; Young, H.A.; Ortaldo, J.R.  
Proc. Natl. Acad. Sci. U.S.A. 90, 542-546, 1993  
A:Title: A cyclophilin-related protein involved in the function of natural killer cells.  
A:Reference number: A47328; MUID:93133824; PMID:8421688  
A:Accession: B47328  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1507 <AND>  
A:CROSS-references: UNIPROT:Q63896; GB:L04289; NID:9192866  
A:Note: authors translated the codon AGT for residue 972 as Arg  
R:Rinfret, A.; Anderson, S.K.  
Mol. Immunol. 30, 1307-1313, 1993  
A:Title: IL-2 regulates the expression of the NK-TR gene via an alternate RNA splicing m  
A:Reference number: 157820; MUID:94019422; PMID:8411330  
A:Accession: 177662  
A:Status: preliminary; translated from GB/EMBL/DBBJ  
A:Molecule type: mRNA  
A:Residues: 235-237:263-294 <RIN>  
A:CROSS-references: GB:S65998; NID:9425701; PIDN:AAB28500.1; PID:9425702  
C:Genetics:  
A:Gene: NK-TR  
C:Superfamily: natural killer cell tumor-recognition protein; cyclophilin homology  
C:Keywords: alternative splicing; lymphocyte  
F:60-230/Domain: cyclophilin homology <CYP>  
Query Match 9.0%; Score 91.5; DB 2; Length 1507;  
Best Local Similarity 22.6%; Pred. No. 15;  
Matches 49; Conservative 24; Mismatches 73; Indels 71; Gaps 10;  
QY 18 PEKIQAFTDDIAKYFSKEWEKMKASEKIFYVMKRYEAMTKLGFKATLPPFMCNKRA 76  
DB 1021 FOKQHSKDDLDGDKTKAREKSKAKK-----DKKHAKPKRKAQAFHWQPPLEFGDDE 1073  
QY 77 EDFQGNLDLNDNPNRGNOVER-----PQMTFGRLOGISPKIMPKKP-----AEEG 120  
DB 1074 EEMNGKQVTQDPKERRHVSEKCEAVKDGIPNVKTECDGSSPS-KPKKGTLEQDPLAEGG 1132  
QY 121 NDSEVP-----BASGP-----QNDGKELCPGPKPTTSEKIHE 153  
DB 1133 HDPSSCPAPLKVEDNTASSPFSQAQLEHFGGSDVLQTDNNMBICTP-----D 1182  
QY 154 RSGPKRGB---HAWTHRLRERKQVLYEISDPEDD 187  
DB 1183 RISPAGKEVVSPLANHRLDSEPVNII-----PEQDE 1213  
RESULT 9  
T21386  
hypoetical protein F29D11.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T21386; T21549  
R:McMurtay, A.  
submitted to the EMBL Data Library, August 1996  
A:Reference number: Z19415  
A:Accession: T21386  
A:Status: preliminary; translated from GB/EMBL/DBBJ  
A:Molecule type: DNA  
A:Residues: 1-1153 <WTL>  
A:CROSS-references: UNIPROT:Q93598; EMBL:Z78419; PIDN:CAB01707.1; GSPDB:GN00019; CESP:F2  
A:Experimental source: clone F28A3  
R:Wilkinson, J.  
submitted to the EMBL Data Library, June 1996  
A:Reference number: Z19439  
A:Accession: T21549  
A:Status: preliminary; translated from GB/EMBL/DBBJ  
A:Molecule type: DNA  
A:Residues: 1-1153 <W12>  
A:CROSS-references: EMBL:Z73907; PIDN:CAA98126.1; GSPDB:GN00019; CESP:F29D11.2

A:Experimental source: clone F29D11  
C:Genetics:  
A:Gene: CESP:F29D11.2  
A:Map position: 1  
A:Introns: 83/2; 117/3; 186/3; 243/3; 278/1; 353/1; 493/2; 543/2; 685/3; 884/2; 944/3; 1  
C:Superfamily: Caenorhabditis elegans hypothetical protein F29D11.2  
Query Match 8.9%; Score 90.5; DB 2; Length 1153;  
Best Local Similarity 25.4%; Pred. No. 13;  
Matches 44; Conservative 22; Mismatches 54; Indels 53; Gaps 10;  
QY 20 KIQAFDDIAKYFSKEWEKMKASEKIFYVMKRYEAMTKLGFKATLPPFMCNKRAEDF 79  
DB 1014 ELRKAHRKIAKILGNEDE-----IEEVPSTDLRKEAAPSKTKRKNANL 1058  
QY 80 QGNLDL--DND-----PNRGNQVERPQMTFGRLOGISPKIMP-----KKPAEAGND 122  
DB 1059 ISTDIADVNDVNMEDDKPGPSRATVVKRP-----APRATPASATKKPLVE-ED 1108  
QY 123 SEEVPEASGPQNDGKELCPGPKPTTSEK---IHERSGPKRGEHAWTHRLRERK 172  
DB 1109 ALEILK-SPPRNTKK---PPSRPTTATRTAATAARTAPPRS----ARKLRSEK 1153  
RESULT 10  
JQ0075  
neuromodulin - goldfish  
N:Alternate names: B-50; F1; GAP-43; pp46  
C:Species: Carassius auratus (goldfish)  
C:Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 09-Jul-2004  
C:Accession: JQ0075  
R:LaBate, M.E.; Skene, J.H.P.  
Neuron 3, 299-310, 1989  
A:Title: Selective conservation of GAP-43 structure in vertebrate evolution.  
A:Reference number: JQ0075; MUID:90380372; PMID:2641999  
A:Accession: JQ0075  
A:Molecule type: mRNA  
A:Residues: 1-213 <LAB>  
A:CROSS-references: UNIPROT:P17691; GB:M26250; NID:9212953; PIDN:AAA03010.1; PID:9212954  
A:Note: this protein contains potential sites for fatty acylation and membrane attachment  
C:Comment: GAP-43 is a major growth cone membrane protein whose expression is widely cor  
C:Comment: Features of the protein are most highly conserved in vertebrate evolution.  
C:Comment: The amino end (1-57) has a strictly conserved domain including protein-protein  
C:Superfamily: histone H1  
C:Keywords: calmodulin binding; lipoprotein; phosphoprotein; thiolester bond  
F:3.4/Binding site: palmitate (Cys) (covalent) #status predicted  
F:81/Binding site: phosphate (Ser) (covalent) #status predicted  
Query Match 8.7%; Score 88.5; DB 2; Length 213;  
Best Local Similarity 23.3%; Pred. No. 2.7;  
Matches 44; Conservative 31; Mismatches 65; Indels 49; Gaps 10;  
QY 19 EKIQAFTDDIAKYFSKEWEKMKASEKIFYVMKRYEAMTKLGFKATLPPFMCNKRAED 78  
DB 12 EKNEEDQGEIKQDGTGKPEENAHKAATKI-----QASFRGHITRKKWKD 54  
QY 79 FQGNLDLND--PNRGNQVERPQMTFGRLOGISPKIMPKKP-----AEEGNDSE--VP 127  
DB 55 -EDKDGENDTAPDESAETEKE-----ERVSPS---EEKPVEVSTETAEESKPAEQPNSP 105  
QY 128 BASGQNDGKELCPGPKPT---TSEKIHRSRGPKRGEHA-----WTHRLRERKQVLYEE 179  
DB 106 AAAPPTAATDSAPSDTPTKEEAQEQLODAEPEKTEENTAADDITTOKEEKE----EE 161  
QY 180 ISDPEEDE 188  
DB 162 EEEEEEE 170  
RESULT 11  
B88734  
protein F32E10.2 [imported] - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans

C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
C/Accession: B88734  
R/anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A>Title: Genome sequence of the nematode C. elegans: a platform for investigating biological processes  
A/Reference number: A75000; MUID:99069613; PMID:9851916  
A/Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_elegans/  
A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A/Accession: B88734  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-270 <STO>  
A/Cross-references: UNIPROT:Q19972; GB:chr\_IV; PIDN:AAA83357.1; PID:gll18175; GSPDB:GN0023  
C/Genetics:  
A/Gene: F3E10.2  
A/Map position: 4

Query Match            8.7%; Score 88.5; DB 2; Length 270;  
Best Local Similarity 24.9%; Pred. No. 3.6;  
Matches 50; Conservative 34; Mismatches 82; Indels 35; Gaps 10;

Qy         3   GDFAFRRPTVGAIPEKIQAFDDIAKYFSKEEWKMKASEKIFYVYMKRYEAMTKLG 62  
            |||||         : : : : : : : : : : : : : : : : : : : : : :  
Db         64   GDSLL-----STYKKVTQSDD-----SSGY-----AVERVL-AHRKVKGSLYLIVQ 107  
            |||||         : : : : : : : : : : : : : : : : : : : : : :

Qy         63   FKATLPPFCNKRAEDFOGNDLDNDPN-----RGNQ-----VERPQMFTGRLOGISPKIM 112  
            |||||         : : : : : : : : : : : : : : : : : : : : : :  
Db         108   WKGYPHVPVNSEWEE---DLNCKDLLAAAYKHQBDELKIAQTTPKKTPTPKTPKSKL 163  
            |||||         : : : : : : : : : : : : : : : : : : : : : :

Qy         113   PKPAEGNDSEEV-PEASQPONDGKELCPGKPTTSEKI--ERSGPGRGEHAWTHRLR 169  
            ::::|         : : : : : : : : : : : : : : : : : : : : : :  
Db         164   KRALTPSDDEESAGPIAPEFKPTKQSTKLRTTSPETNLVEKSKKAIPDLENHTLD 223  
            ::::|         : : : : : : : : : : : : : : : : : : : : : :

Qy         170   ERKQLVI--YEESDPDEEDE 188  
            : : : : : : : : : : : : : : : : : : : : : :  
Db         224   QEKNDVIERVEIQEDEDDED 244  
            : : : : : : : : : : : : : : : : : : : : : :

RESULT 12  
T23908  
hypothetical protein R04F11.3 - Caenorhabditis elegans  
C/Species: Caenorhabditis elegans  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C/Accession: T23908  
R/Harris, B.  
Submitted to the EMBL Data Library, June 1996  
A/Reference number: Z19816  
A/Accession: T23908  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-445 <WIL>  
A/Cross-references: UNIPROT:Q21710; EMBL:Z74475; PIDN:CAA98958.1; GSPDB:GN00023; CESP:R04F11  
A/Experimental source: Clone R04F11  
C/Genetics:  
A/Gene: CESP:R04F11.3  
A/Map position: 5  
A/Introns: 60/3; 94/2; 396/1

Query Match            8.7%; Score 88; DB 2; Length 445;  
Best Local Similarity 21.0%; Pred. No. 7.1;  
Matches 43; Conservative 36; Mismatches 60; Indels 66; Gaps 8;

Qy         14   GAQIPEKI--OKAFDDIAKYFSKEEWKMKASEKIFYVYMKRYEAMTKLGFKATLPPFM 71  
            | : : : : : : : : : : : : : : : : : : : : : :  
Db         183   GKXLPQLRPQKSDS-----SNVKVEKSTEQ-----KPTIEPST 219  
            | : : : : : : : : : : : : : : : : : : : : : :

Qy         72   CNKRAEDFOGNDLDNDPNRG-----NQVERPQMFTGRLOGISPKIMPKPAAEEN 121  
            : : : : : : : : : : : : : : : : : : : : : :  
Db         220   SSKSALETKKSIAPPKKQGTLSMFSKTEKTEKTSQSPVQDQEKKNKRPMIEWD 279  
            : : : : : : : : : : : : : : : : : : : : : :  
Qy         122   DSEEVPEASQPONDGKELCPGKPTTSEKI-----HERSGPKRGEHA 163  
            : : : : : : : : : : : : : : : : : : : : : :  
Db         280   ESEELIKATKP-----KEKTEKPP---EKLIISODSEFFSSREOSVEKESPPAKRKRV 332  
            : : : : : : : : : : : : : : : : : : : : : :

A;Gene: GDB:NHCP2  
A;Cross-references: GDB:119451; OMIM:118880  
A;Map position: 7pter-7qter  
C;Superfamily: nonhistone chromosomal protein HMG-2; HMG box homology  
C;Keywords: chromosomal protein; DNA binding; nucleus  
F:6-83/Domain: HMG box homology <HMG1>  
F:92-166/Domain: HMG box homology <HMG2>

Query Match 8.6%; Score 87.5; DB 1; Length 209;  
Best Local Similarity 24.2%; Pred. No. 3.2; Mismatches 66; Indels 65; Gaps 8;  
Matches 47; Conservative 16;

Qy 32 FSK---EWEKAKASEKIFYYVMKRYEAMTKLGFKATLPPFMCNKRAEDFQGNLDLND 87  
Db 41 FSKCSERWKTSAKESKPEDMAKSDKARYDREMKNVPP-----KGDKKKKKDPA 94

Qy 88 PNRGNQVERPQMTGRLQGISPKMPKPA-----EKGNDSEVP----- 127  
Db 95 PKR-----PPSAFFLFCSEHRPKIKSEHPGLSIGDTAKKLGEMWSEQSAKDQPYEQKAA 149

Qy 128 -----EASGPQNDGKELCPGKPTTSEKIHRSQKGEHAWTHRLRERKQL 174  
Db 150 KLKKEKYEKDIAYRAKGESEAGK--GECRTGSKKNEPEDEEEEEE----- 195

Qy 175 VIYEEISDPEEDE 188  
Db 196 ---EEDEDEEEDE 206

RESULT 15  
T48836  
hypothetical protein 68B2.200 [imported] - Neurospora crassa  
C;Species: Neurospora crassa  
C;Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 09-Jul-2004  
C;Accession: T48836  
R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Partmann, B.; Holland, R.; Nyakatura, submitted to the Protein Sequence Database, April 2000  
A;Reference number: Z24541  
A;Accession: T48836  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-541 <SCH>  
A;Cross-references: UNIPROT:Q9PGV6; EMBL:AL353821; GSPDB:GN00112; NCSP:68B2.200  
A;Experimental source: cosmid contig 68B2; strain 74  
C;Genetics:  
A;Gene: NCSP:68B2.200  
A;Map position: 2  
A;Introns: 35/3

Query Match 8.6%; Score 87.5; DB 2; Length 541;  
Best Local Similarity 23.2%; Pred. No. 9.8;  
Matches 33; Conservative 23; Mismatches 57; Indels 29; Gaps 7;

Qy 18 PEKIQKAFDDIAKFSKEWEKAKASEKIFYYVMKRYEAMTKLGF-KATLPPFMCNKRA 76  
Db 362 PE--QQLLPTVAKRLQEKWER---EGKVGNVYDK-EFRPLTDEGFLEPPPEHPIVAPEHE 415

Qy 77 EDFQGNLDNDPNRGNQVERPQMTGRLQGISPKMPKPAEEGNDSEVPESASGPQNDG 136  
Db 416 PAENENENENENENENEKEEQ-----EKENEKEQ-PQADWPLKPA 455

Qy 137 KELCPPGKPTTSEKIHRSQK 158  
Db 456 EPQCPQSPPPPS--LHSLQSPR 475

Search completed: June 20, 2005, 12:22:05  
Job time : 42 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2005, 12:03:35 ; Search time 174 Seconds  
(without alignments)  
553.281 Million cell updates/sec

Title: US-10-777-053-40

Perfect score: 1012

Sequence: 1 MNGDDAFARRFTVGAQIPEK.....RRKQLVIYEISDPEEDDE 188

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	ID	Description
1	1012	100.0	1 SSX2_HUMAN	Q16385 homo sapien
2	905	89.4	1 SSX3_HUMAN	Q99909 homo sapien
3	876	86.6	2 Q7RTT3	Q7rtt3 homo sapien
4	852	84.2	2 Q7RTT5	Q7rtt5 homo sapien
5	847	83.7	1 SSX5_HUMAN	O60225 homo sapien
6	845.5	83.5	2 Q9BU88	O9bu88 homo sapien
7	798	78.9	1 SSX4_HUMAN	O60224 homo sapien
8	791	78.2	2 Q7RTT6	Q7rtt6 homo sapien
9	772	76.3	1 SSX1_HUMAN	Q16384 homo sapien
10	735	72.6	2 Q9BRW7	Q9brw7 homo sapien
11	618	61.1	2 Q96QI1	Q96qi1 homo sapien
12	587.5	58.1	2 Q96QI0	Q96qi0 homo sapien
13	567	56.0	2 Q7RTT4	Q7rtt4 homo sapien
14	427.5	42.2	2 Q9Y444	Q9y444 homo sapien
15	335	33.1	4 Q8WWZ9	Q8wwz9 homo sapien
16	258.5	25.5	1 Q9NZK4	Q9nz44 homo sapien
17	253	25.0	2 Q6XAR4	Q6xar4 mus musculus
18	238.5	23.6	2 Q6XAR8	Q6xar8 mus musculus
19	229.5	22.7	2 Q6XAR7	Q6xar7 mus musculus
20	227.5	22.5	2 Q6XAS4	Q6xas4 mus musculus
21	225.5	22.3	2 Q8OZT4	Q8ozt4 mus musculus
22	225.5	22.3	2 Q8C5Z3	Q8c5z3 mus musculus
23	222.5	22.0	2 Q6XAS2	Q6xas2 mus musculus
24	215.5	21.3	2 Q6XAR6	Q6xar6 mus musculus
25	215.5	21.3	2 Q9CFU1	Q9cfu1 m mus muscu
26	211.5	20.9	2 Q6XAR5	Q6xar5 mus musculus
27	206.5	20.4	2 Q6XAS3	Q6xas3 mus musculus
28	192.5	19.0	2 Q7S101	Q7s101 homo sapien
29	163	16.1	39 Q81ZG5	Q81zg5 homo sapien
30	163	16.1	39 Q81ZG6	Q81zg6 homo sapien
31	163	16.1	39 Q81ZG7	Q81zg7 homo sapien

32	163	16.1	39	2 Q81ZG8	Q81zg8 homo sapien
33	163	16.1	39	2 Q81ZG9	Q81zg9 homo sapien
34	163	16.1	39	2 Q81ZH0	Q81zh0 homo sapien
35	149.5	14.8	140	2 Q6XAR9	Q6xar9 mus musculus
36	149	14.7	140	2 Q6XAS1	Q6xas1 mus musculus
37	148	14.6	140	2 Q6XAS0	Q6xas0 mus musculus
38	135	13.3	752	2 Q61V72	Q61v72 homo sapien
39	124	12.3	281	2 Q6ZMS7	Q6zms7 homo sapien
40	121	12.0	39	2 Q81ZH1	Q81zh1 homo sapien
41	119	11.8	524	2 Q8N2J5	Q8n2j5 homo sapien
42	119	11.8	785	2 Q9ULD5	Q9uld5 homo sapien
43	118	11.7	276	2 Q8WV14	Q8wv14 homo sapien
44	118	11.7	782	2 Q7S701	Q7s701 homo sapien
45	115	11.4	642	1 Z398_HUMAN	Q8tcl7 homo sapien

#### ALIGNMENTS

RESULT 1  
SSX2\_HUMAN  
ID SSX2\_HUMAN STANDARD; PRT; 188 AA.  
AC Q16385; Q16404; Q96IP7;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE SSX2 protein (synovial sarcoma, X breakpoint 2) (SSX) (HOM-MEL-40).  
GN Name=SSX2;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
RX MEDLINE=95292374; PubMed=7539744;  
RA Crew A.J., Clark J., Fisher C., Gill S., Grimer R., Chand A.,  
RA Shipley J., Gusterson B.A., Cooper C.S.;  
RT "Fusion of SYT to two genes, SSX1 and SSX2, encoding proteins with  
RT homology to the Kruppel-associated box in human synovial sarcoma.";  
RL EMBO J. 14:2333-2340(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Skin;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmitz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.L., Skalska U., Smallos D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [3]  
RP SEQUENCE OF 111-188 FROM N.A.  
RC TISSUE=Synovial sarcoma;  
RX MEDLINE=95384157; PubMed=7655467;  
RA de Leeuw B., Balemans M., Olde weghuis D., Geurts van Kessel A.;  
RT "Identification of two alternative fusion genes, SYT-SSX1 and SYT-  
RT SSX2, in t(X;18) (p11.2;q11.2)-positive synovial sarcomas.";  
RL Hum. Mol. Genet. 4:1097-1099(1995).

[4]  
RP SEQUENCE OF 111-188 FROM N.A. (SSXT-SSX2 FUSION PROTEIN).  
RC TISSUE-Synovial sarcoma;  
RX MEDLINE=95038836; PubMed=7951320;  
RA Clark J., Rocques P.J., Crew A.J., Gill S., Shipley J., Chan A.M.-L.,  
RA Gusterson B.A., Cooper C.S.;  
RT "Identification of novel genes, SYT and SSX, involved in the  
RT t(X;18)(p11.2;q11.2) translocation found in human synovial sarcoma.";  
RL Nat. Genet. 7:502-508(1994).  
RN [5]  
RP SEQUENCE OF 68-116 FROM N.A. (SSXT-SSX2 FUSION PROTEIN).  
RX MEDLINE=96094743; PubMed=7495284;  
RA Fligman I., Lonardo F., Jhanwar S.C., Gerald W.L., Woodruff J.,  
RA Ladanyi M.;  
RT "Molecular diagnosis of synovial sarcoma and characterization of a  
RT variant SYT-SSX2 fusion transcript.";  
RL Am. J. Pathol. 147:1592-1599(1995).  
RN [6]  
RP INTERACTION WITH SSX2IP.  
RX PubMed=12007189; DOI=10.1002/gcc.10073;  
RA de Bruijn D.R.H., dos Santos N.R., Kater-Baats E., Thijssen J.,  
RA van den Berk L., Stap J., Bailemans M., Schepens M., Merx G.,  
RA van Kessel A.G.;  
RT "The cancer-related protein SSX2 interacts with the human homologue of  
RT a Ras-like GTPase interactor, RAB3IP, and a novel nuclear protein,  
RT SSX2IP.";  
RL Genes Chromosomes Cancer 34:285-298(2002).  
CC -!- FUNCTION: Could act as a modulator of transcription.  
CC -!- SUBUNIT: Interacts with SSX2IP.  
CC -!- TISSUE SPECIFICITY: Expressed at high level in the testis.  
CC Expressed at low level in thyroid. Not detected in tonsil, colon,  
CC lung, spleen, prostate, kidney, striated and smooth muscles.  
CC Detected in rhabdomyosarcoma and fibrosarcoma cell lines. Not  
CC detected in mesenchymal and epithelial cell lines.  
CC -!- DISEASES: Participates in a t(X;18)(p11.2;q11.2) chromosomal  
CC translocation that is specifically found in more than 80% of  
CC synovial sarcoma and produces the SSXT-SSX1 or SSXT-SSX2 fusion  
CC products. These hybrid proteins are probably responsible for  
CC transforming activity. Heterogeneity in the position of the  
CC breakpoint can occur (low frequency).  
CC -!- SIMILARITY: Belongs to the SSX family.  
CC -!- SIMILARITY: Contains 1 KRAB-related domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X86175; CAA60111.1; -;  
DR EMBL; BC007343; AAH07343.1; -;  
DR EMBL; BC016957; AAH16957.1; -;  
DR EMBL; S79332; AAB35379.1; -;  
DR EMBL; X79200; -; NOT ANNOTATED CDS.  
DR EMBL; S79894; AAB35674.1; ALT\_INIT.  
DR PIR; S55058; S55058  
DR Genew; HGNC:11336; SSX2.  
DR H-InvDB; HIX0016747; -;  
DR MIM; 300192; -;  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.  
DR InterPro; IPR001909; KRAB.  
DR InterPro; IPR003655; KRAB-related.  
DR Pfam; PF01352; KRAB; 1.  
DR SMART; SM00349; KRAB; 1.  
DR PROSITE; PS0806; KRAB RELATED; 1.  
KW Chromosomal translocation; Multigene family; Proto-oncogene;  
KW Transcription regulation.  
FT DOMAIN 20 83 KRAB-related.  
FT SITE 68 69 Breakpoint for translocation to form the  
FT SITE 110 111 SSXT-SSX2 fusion protein (rare).  
FT Breakpoint for translocation to form the

FT SSXT-SSX2 fusion protein.  
FT R -> P (in Ref. 2; AAH07343).  
SQ SEQUENCE 169 169  
SEQUENCE 188 AA; 21620 MW; BF5D18AA5F45B1B1 CRC64;  
Query Match 100.0%; Score 1012; DB 1; Length 188;  
Best local Similarity 100.0%; Pred. No. 1.3e-72;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MNGDDAFARRPTVGAQIPEKIQAFFDDIAKYFSKEWEKMKASEKIFVYMKRYEAMTK 60  
Db 1 MNGDDAFARRPTVGAQIPEKIQAFFDDIAKYFSKEWEKMKASEKIFVYMKRYEAMTK 60  
Qy 61 LGFKATLPPPMCNKRAEDFQGNLDNDPNRGNQVERPQMTFGRLOGISPKIMPKPAEEG 120  
Db 61 LGFKATLPPPMCNKRAEDFQGNLDNDPNRGNQVERPQMTFGRLOGISPKIMPKPAEEG 120  
Qy 121 NDSEVPEASGPONDGKELCPGKPTTSEKIHSSGPKRGEHAWTHRLRRKQLVYEEI 180  
Db 121 NDSEVPEASGPONDGKELCPGKPTTSEKIHSSGPKRGEHAWTHRLRRKQLVYEEI 180  
Qy 181 SDPEEDDE 188  
Db 181 SDPEEDDE 188  
RESULT 2  
SSX3 HUMAN  
ID -SSX3 HUMAN STANDARD; PRT; 188 AA.  
AC Q99909; O60223;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE SSX3 protein.  
GN Name=SSX3;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
ON NCBI\_TaxID=9606;  
RX [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Fibrosarcoma, and Testis;  
RX MEDLINE=96302330; PubMed=8697803;  
RA de Leeuw B., Bailemans M., Geurts van Kessel A.;  
RT "A novel Kruppel-associated box containing the SSX gene (SSX3) on the  
RT human X chromosome is not implicated in t(X;18)-positive synovial  
RT sarcomas.";  
RL Cytoogenet. Cell Genet. 73:179-183(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98021352; PubMed=9378559;  
RA Gure A.O., Tuereci O., Sahin U., Tsang S., Scanlan M.J., Jager E.,  
RA Knuth A., Pfeundschoh M., Old L.J., Chen Y.-T.;  
RT "SSX: a multigene family with several members transcribed in normal  
RT testis and human cancer.";  
RL Int. J. Cancer 72:965-971(1997).  
RN [3]  
RP INTERACTION WITH SSX2IP.  
RX PubMed=12007189; DOI=10.1002/gcc.10073;  
RA de Bruijn D.R.H., dos Santos N.R., Kater-Baats E., Thijssen J.,  
RA van den Berk L., Stap J., Bailemans M., Schepens M., Merx G.,  
RA van Kessel A.G.;  
RT "The cancer-related protein SSX2 interacts with the human homologue of  
RT a Ras-like GTPase interactor, RAB3IP, and a novel nuclear protein,  
RT SSX2IP.";  
RL Genes Chromosomes Cancer 34:285-298(2002).  
CC -!- FUNCTION: Could act as a modulator of transcription.  
CC -!- SUBUNIT: Interacts with SSX2IP.  
CC -!- SIMILARITY: Belongs to the SSX family.  
CC -!- SIMILARITY: Contains 1 KRAB-related domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X86175; CAA60111.1; -;  
DR EMBL; BC007343; AAH07343.1; -;  
DR EMBL; BC016957; AAH16957.1; -;  
DR EMBL; S79332; AAB35379.1; -;  
DR EMBL; X79200; -; NOT ANNOTATED CDS.  
DR EMBL; S79894; AAB35674.1; ALT\_INIT.  
DR PIR; S55058; S55058  
DR Genew; HGNC:11336; SSX2.  
DR H-InvDB; HIX0016747; -;  
DR MIM; 300192; -;  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.  
DR InterPro; IPR001909; KRAB.  
DR InterPro; IPR003655; KRAB-related.  
DR Pfam; PF01352; KRAB; 1.  
DR SMART; SM00349; KRAB; 1.  
DR PROSITE; PS0806; KRAB RELATED; 1.  
KW Chromosomal translocation; Multigene family; Proto-oncogene;  
KW Transcription regulation.  
FT DOMAIN 20 83 KRAB-related.  
FT SITE 68 69 Breakpoint for translocation to form the  
FT SITE 110 111 SSXT-SSX2 fusion protein (rare).  
FT Breakpoint for translocation to form the



use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; S82471; AAC37436.2; -;  
DR EMBL; U0840; AAC05819.1; -;  
DR Genew; HGNC:11337; SSX3.  
DR MIM; 300325; -;  
DR InterPro; IPR001909; KRAB.  
DR InterPro; IPR003655; KRAB\_related.  
DR Pfam; PF01352; KRAB; 1.  
DR SMART; SM00349; KRAB; 1.  
DR PROSITE; PS50806; KRAB\_RELATED; 1.  
KW Multigene family; Transcription regulation.  
FT DOMAIN 20 83 KRAB-related.  
FT CONFLICT 95 L -> Q (in Ref. 2).  
SQ SEQUENCE 188 AA; 21682 MW; 368C7487C1450E72 CRC64;

Query Match 89.4%; Score 905; DB 1; Length 188;  
Best Local Similarity 91.0%; Pred. No. 4.1e-64;  
Matches 171; Conservative 3; Mismatches 14; Indels 0; Gaps 0;  
QY 1 MNGDDAFARRPTVGAQIPEKIQAFDDIAKYFSKEWEKMKASEKIFVYVMKRYEAMTK 60  
DB 1 MNGDDTFARRPTVGAQIPEKIQAFDDIAKYFSKEWEKMKYSEKIVYVMKRYEAMTK 60  
QY 61 LGFKATLPPFMCMKRAEDFQGNLDNDPNRGNQVVERPQMTFGRLOGISPKIMPKPAEAG 120  
DB 61 LGFKAILPSFMENKRVTFQGNDFDNDPNRGNQVLRPQMTFGRLOGIFPKIMPKPAEAG 120  
QY 121 NDSEVPEASGPNQDGKELCPGKPTTSEKIHRSQKGEHAWTHRLRERKOLVYEEI 180  
DB 121 NVSKEVPEASGPNQDGKQLCPPGKPTTSEKINMISGPKRGEHAWTHRLRERKOLVYEEI 180  
QY 181 SDPEEDDE 188  
DB 181 SDPEEDDE 188

## RESULT 3

QYRTT3  
ID QYRTT3 PRELIMINARY; PRT; 188 AA.  
AC QYRTT3;  
DT 01-MAR-2004 (TremBLrel. 26, Created)  
DT 01-MAR-2004 (TremBLrel. 26, Last sequence update)  
DE Synovial sarcoma X breakpoint 9 protein.  
GN Name-SSX9;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22204129; PubMed=12216073;  
RA Gure A.O., Wei I.J., Old L.J., Chen Y.-T.;  
RT "The SSX gene family: characterization of 9 complete genes.";  
RL Int. J. Cancer 101:448-453(2002).  
CC -!- MISCELLANEOUS: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ third party annotation (TFA) entry.  
DR EMBL; BK000689; DAA00376.1; -;  
DR GO; GO:0005622; C:intracellular; IEA.  
DR GO; GO:0003676; F:nucleic acid binding; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR InterPro; IPR001909; KRAB\_related.  
DR InterPro; IPR003655; KRAB.  
DR Pfam; PF01352; KRAB; 1.  
DR PROSITE; PS50806; KRAB\_RELATED; 1.  
SQ SEQUENCE 188 AA; 21553 MW; 4DC6C4FBA282CEFC CRC64;

Query Match 86.6%; Score 876; DB 2; Length 188;  
Best Local Similarity 87.2%; Pred. No. 8.2e-62;

Matches 164; Conservative 8; Mismatches 16; Indels 0; Gaps 0;  
QY 1 MNGDDAFARRPTVGAQIPEKIQAFDDIAKYFSKEWEKMKASEKIFVYVMKRYEAMTK 60  
DB 1 MNGDDAFARRPRAGSIPEKIQAFDDIAKYFSKEWEKMKSEKIIYYVMKRYEAMTK 60  
QY 61 LGFKATLPPFMCMKRAEDFQGNLDNDPNRGNQVVERPQMTFGRLOGISPKIMPKPAEAG 120  
DB 61 LGFKATLPPFMCMKRAEDFQGNLDNDPNRGNQVVERPQMTFGRLOGIFPKIMPKPAEAG 120  
QY 121 NDSEVPEASGPNQDGKELCPGKPTTSEKIHRSQKGEHAWTHRLRERKOLVYEEI 180  
DB 121 NDSKEVPEASGPNQDGKQLCPPGKPTTSEKINKASGPKRGEHAWTHRLRERKOLVYEEI 180  
QY 181 SDPEEDDE 188  
DB 181 SDPEEDDE 188

## RESULT 4

QYRTT5  
ID QYRTT5 PRELIMINARY; PRT; 188 AA.  
AC QYRTT5;  
DT 01-MAR-2004 (TremBLrel. 26, Created)  
DT 01-MAR-2004 (TremBLrel. 26, Last sequence update)  
DE Synovial sarcoma X breakpoint 7 protein.  
GN Name-SSX7;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22204129; PubMed=12216073;  
RA Gure A.O., Wei I.J., Old L.J., Chen Y.-T.;  
RT "The SSX gene family: characterization of 9 complete genes.";  
RL Int. J. Cancer 101:448-453(2002).  
CC -!- MISCELLANEOUS: The sequence shown here is derived from an  
CC EMBL/GenBank/DBJ third party annotation (TFA) entry.  
DR EMBL; BK000687; DAA00374.1; -;  
DR GO; GO:0005622; C:intracellular; IEA.  
DR GO; GO:0003676; F:nucleic acid binding; IEA.  
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
DR InterPro; IPR001909; KRAB\_related.  
DR InterPro; IPR003655; KRAB.  
DR Pfam; PF01352; KRAB; 1.  
DR PROSITE; PS50806; KRAB\_RELATED; 1.  
SQ SEQUENCE 188 AA; 21591 MW; 3871D4FAA13DC78E CRC64;

Query Match 84.2%; Score 852; DB 2; Length 188;  
Best Local Similarity 85.6%; Pred. No. 6.6e-60;  
Matches 161; Conservative 9; Mismatches 18; Indels 0; Gaps 0;

QY 1 MNGDDAFARRPTVGAQIPEKIQAFDDIAKYFSKEWEKMKASEKIFVYVMKRYEAMTK 60  
DB 1 MNGDDAFARRPRAGSIPEKIQAFDDIAKYFSKEWEKMKSEKISYYVMKRYEAMTK 60  
QY 61 LGFKATLPPFMCMKRAEDFQGNLDNDPNRGNQVVERPQMTFGRLOGISPKIMPKPAEAG 120  
DB 61 LGFKATLPPFMCMKRAEDFQGNLDNDPNRGNQVVERPQMTFGRLOGIFPKIMPKPAEAG 120  
QY 121 NDSEVPEASGPNQDGKELCPGKPTTSEKIHRSQKGEHAWTHRLRERKOLVYEEI 180  
DB 121 NDSKGVPEASGPNQDGKHLCPGKPTTSEKINKTSKPKRGEHAWTHRLRERKOLVYEEI 180  
QY 181 SDPEEDDE 188  
DB 181 SDPEEDDE 188

RESULT 5  
SSX5\_HUMAN

```
ID AC 060225; Q96AWJ; STANDARD; PRT; 188 AA.
DC 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE SSX5 protein.
GN Name=SSX5;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=98021352; PubMed=9378559;
RA Gure A.O., Tuercil O., Sahin U., Tsang S., Scanlan M.J., Jager E.,
RA Knuth A., Pfeundschnig M., Old L.J., Chen Y.-T.;
RT "SSX: a multigene family with several members transcribed in normal
RT testis and human cancer.";
RL Int. J. Cancer 72:965-971(1997).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Skin;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., McKernan K.J., Abramson R.D., Mullahy S.J.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Could act as a modulator of transcription.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=O60225-1; Sequence=Displayed;
CC Name=2;
CC IsoId=O60225-2; Sequence=VSP_006274;
CC Note=No experimental confirmation available;
CC -!- SIMILARITY: Belongs to the SSX family.
CC -!- SIMILARITY: Contains 1 KRAB-related domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL; U90842; AAC05021.1; -
CC EMBL; BC016640; AAH16640.1; -
CC Genew; HGNC:11339; SSX5.
CC MIM; 300327; -
CC InterPro; IPR001909; KRAB.
CC InterPro; IPR003655; KRAB-related.
CC Pfam; PF01352; KRAB; 1.
CC SMART; SM00349; KRAB; 1.
CC PROSITE; PS0806; KRAB RELATED; 1.
CC Alternative splicing; Multigene family; Transcription regulation.
CC DOMAIN 20 83 KRAB-related.
```

```
FT VARSPLIC 23 23 K -> KHPWQVCDRGHVLNLSFPWVKVGRBPASSIKALLC
FT GRGEAR (in isoform 2).
FT /FTid=VSP_006274.
FT CONFLICT 184 184 P -> O (in Ref. 2).
SQ SEQUENCE 188 AA; 21628 MW; AD2A3096931C5E37 CRC64;
Query Match 83.7%; Score 847; DB 1; Length 188;
Best Local Similarity 83.5%; Pred. No. 1.6e-59;
Matches 157; Conservative 14; Mismatches 17; Indels 0; Gaps 0;
Qy 1 MNGDDAFARRPTVGAQIPEKIQAFDDIAKYFSKEWEKMKASEKIFVYMKRYEAMTK 60
Db 1 MNGDDAFVRRPRVGSQIPQKQKAFDDIAKYFSKEWEKMKASEKIIIVYMKRYEAMTK 60
Qy 61 LGFKATLPPPFMCNKRADFQGNLDLNDPNRGNQVERPQMTFGRLOGISPKIMPKPAEAG 120
Db 61 LGFKATLPPFWRNKRVDQGNDFDNDPNRGNQVEHQMTFGRLOGIFPKITPEKPAEAG 120
Qy 121 NDSKEVPEASGFQNDGKELCPGKPTTSEKITHRSKPKRGSHAWTHRLRERKQLVIYBEI 180
Db 121 NDSKGVPEASGFQNNKQLRSPGKLTSEKYNKTSKPKRGKSHAWTHRVRRERKQLVIYBEI 180
Qy 181 SDPEEDDE 188
Db 181 SDPPEDDE 188
RESULT 6
Q9BU88 PRELIMINARY; PRT; 223 AA.
ID Q9BU88
AC Q9BU88;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Synovial sarcoma, X breakpoint 2, isoform a.
GN Name=SSX2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta, and Skin;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E., Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RN Submitted (FEB-2001) to the EMBL/GenBank/DBSJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
```

```

RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC002818; AA02818.1; -.
DR EMBL; BC071827; AA071827.1; -.
DR GO; GO:0003622; C:intracellular; IEA.
DR GO; GO:0003676; P:nucleic acid binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR003655; KRAB-related.
DR Pfam; PF01352; KRAB; 1.
DR SMART; SM00349; KRAB; 1.
DR PROSITE; PS50806; KRAB RELATED; 1.
SQ SEQUENCE 223 AA; 25173 MW; 2BF8E1FFA4D58094 CRC64;

Query Match
Best Local Similarity 83.5%; Score 845.5; DB 2; Length 223;
Matches 160; Conservative 3; Mismatches 5; Indels 13; Gaps 2;

QY 1 MNGDDAFARRPTVGAQIPEIKIQAEDDIKAFDDIAKYFSKEWEKMKASEKIFVYMKRYEAMTK 60
Db 1 MNGDDAFARRPTVGAQIPEIKIQAEDDIKAFDDIAKYFSKEWEKMKASEKIFVYMKRYEAMTK 60
QY 61 LGFKATLPPFMCNKRAEDFQGNLDNDPNRGNQVERPQMTFGRLOGISPKIMPKPAERG 120
Db 61 LGFKATLPPFMCNKRAEDFQGNLDNDPNRGNQVERPQMTFGRLOGISPKIMPKPAERG 120
QY 121 NDSEVPEASGPNQDKELCPGKPTTSEKIHRSQPKGE-----HAW-----THR 167
Db 121 NDSEVPEASGPNQDKELCPGKPTTSEKIHRSQPKGE-----HAW-----THR 167
QY 168 L 168
Db 181 I 181

RESULT 7
SSX4 HUMAN STANDARD; PRT; 188 AA.
AC 060224; Q9UJU9; -.
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE SSX4 protein.
GN Name=SSX4;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98021352; PubMed=9378559;
RA Gure A.O., Tuercio O., Sahin U., Tsang S., Scanlan M.J., Jager E.,
RA Knuth A., Pfeunderschuh M., Old L.J., Chen Y.-T.;
RT "SSX: a multigene family with several members transcribed in normal
RT testis and human cancer.";
RL Int. J. Cancer 72:965-971(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Urinary bladder;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Murzyn D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

```

```

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra W.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE OF 1-155 FROM N.A.
RA Blechschmidt K., Nyakatura G., Strom T.M., Drescher B., Menzel U.,
RA Meindl A., Rosenthal A.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Could act as a modulator of transcription.
CC -!- SIMILARITY: Belongs to the SSX family.
CC -!- SIMILARITY: Contains 1 KRAB-related domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U00841; AAC05820.1; -.
DR EMBL; BC005325; AA05325.1; -.
DR EMBL; AF196972; AAF06796.1; -.
DR Genew; HGNC:11338; SSX4.
DR H-InvDB; HIX0016772; -.
DR MIM; 300326; -.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR003655; KRAB-related.
DR Pfam; PF01352; KRAB; 1.
DR SMART; SM00349; KRAB; 1.
DR PROSITE; PS50806; KRAB RELATED; 1.
KW Multigene family; Transcription regulation.
FT DOMAIN 20 83 KRAB-related.
SQ SEQUENCE 188 AA; 21858 MW; 4ACA2A8737507AE5 CRC64;

Query Match 78.9%; Score 798; DB 1; Length 188;
Best Local Similarity 79.8%; Pred. No. 1.3e-55;
Matches 150; Conservative 12; Mismatches 26; Indels 0; Gaps 0;

QY 1 MNGDDAFARRPTVGAQIPEIKIQAEDDIKAFDDIAKYFSKEWEKMKASEKIFVYMKRYEAMTK 60
Db 1 MNGDDAFARRPTVGAQIPEIKIQAEDDIKAFDDIAKYFSKEWEKMKASEKIFVYMKRYEAMTK 60
QY 61 LGFKATLPPFMCNKRAEDFQGNLDNDPNRGNQVERPQMTFGRLOGISPKIMPKPAERG 120
Db 61 LGFKATLPPFMCNKRAEDFQGNLDNDPNRGNQVERPQMTFGRLOGISPKIMPKPAERG 120
QY 121 NDSEVPEASGPNQDKELCPGKPTTSEKIHRSQPKGEHAWTHRLRERKQLYVEI 180
Db 121 NGLKEVPEASGPNQDKELCPGPNPSTLEKINKTSGPKRGKIAWTHRLRERKQLYVEI 180
QY 181 SDPEEDDE 188
Db 181 SDPEEDDE 188

RESULT 8
Q7RTT6 PRELIMINARY; PRT; 188 AA.
AC Q7RTT6;
DT 01-MAR-2004 (TReMBLrel. 26, Created)
DT 01-MAR-2004 (TReMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Synovial sarcoma X breakpoint 6 protein.
GN Name=SSX6;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

```

```
RP SEQUENCE FROM N.A.
RX MEDLINE=22204129; PubMed=12216073;
RA Gure'A.O., Wei I.J., Old L.J., Chen Y.-T.;
RT "The SSX gene family: characterization of 9 complete genes.";
RL Int. J. Cancer 101:448-453(2002).
CC -!- MISCELLANEOUS: the sequence shown here is derived from an
CC -!- EMBL/GenBank/DBJ third party annotation (TPA) entry.
DR EMBL; BK000686; DAA00373.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR003655; KRAB_related.
DR Pfam; PF01352; KRAB; 1.
DR PROSITE; PS50806; KRAB_RELATED; 1.
SQ SEQUENCE 188 AA; 21688 MW; BBAB5A4F7BA879C8 CRC64;

Query Match 78.2%; Score 791; DB 2; Length 188;
Best Local Similarity 80.9%; Pred. No. 4.6e-55;
Matches 152; Conservative 12; Mismatches 24; Indels 0; Gaps 0;

Qy 1 MNGDDAFARRPTVGAQIPEKIQKAFDDIAKYFSKEEWMKASEKIFVYMKRYEAMTK 60
Db 1 MNGDDAFARRPRDDAKASEKSKAFDDIAKYFSKEEWMKPFSEKISCVMKRYEAMTK 60

Qy 61 LGFKATLPFPFCKRAEDFQGNLDNDNRGNQVERPQMTFGRLOGISPKIMPKPAEEG 120
Db 61 LGFNVTLSLFWRNKXATDSQRNDSNDRNGNEVERPQMTFGRLORIIPKIMPKPAEEG 120

Qy 121 NDSEVVPASGPQNDGKCLPPGKPTTSEKIHRSRGPGRGHAWTHRLRERKQLVYEEI 180
Db 121 SDSKGVPEASGPQNDGKCLPPGKASSEKIHRSRGPGRGHAWTHRLRERKQLVYEEI 180

Qy 181 SDPEEDDE 188
Db 181 SDPEEDDK 188

RESULT 9
SSX1_HUMAN STANDARD; PRT; 188 AA.
ID SQX1_HUMAN
AC Q16384;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE SSX1 protein (Synovial sarcoma, X breakpoint 1).
GN Name=SSX1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Synovial sarcoma;
RX MEDLINE=95292974; PubMed=7539744;
RA Crew A.J., Clark J., Fisher C., Gill S., Grimer R., Chand A.,
RA Shipley J., Gusterson B.A., Cooper C.S.;
RT "Fusion of SYT to two genes, SSX1 and SSX2, encoding proteins with
RT homology to the Kruppel-associated box in human synovial sarcoma.";
RL EMBO J. 14:2333-2340(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=22389257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.F.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullihy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Marra M.A.;
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE OF 111-188 FROM N.A.
RC TISSUE=Synovial sarcoma;
RX MEDLINE=95384157; PubMed=7655467;
RA de Leeuw B., Bailemans M., Olde Weghuis D., Geurts van Kessel A.;
RT "Identification of two alternative fusion genes, SYT-SSX1 and SYT-
RT SSX2, in t(X;18)(p11.2;q11.2)-positive synovial sarcomas.";
RL Hum. Mol. Genet. 4:1097-1099(1995).
CC -!- FUNCTION: Could act as a modulator of transcription.
CC -!- TISSUE SPECIFICITY: Expressed at high level in the testis.
CC Expressed at low level in thyroid. Not detected in tonsil, colon,
CC lung, spleen, prostate, kidney, striated and smooth muscles.
CC Detected in rhabdomyosarcoma and fibrosarcoma cell lines. Not
CC detected in mesenchymal and epithelial cell lines.
CC -!- DISEASE: Participates in a t(X;18)(p11.2;q11.2) chromosomal
CC translocation that is specifically found in more than 80% of
CC synovial sarcoma and produces the SSX1-SSX2 or SSX2-SSX1 fusion
CC products. These hybrid proteins are probably responsible for
CC transforming activity. Heterogeneity in the position of the
CC breakpoint can occur (low frequency).
CC -!- SIMILARITY: Belongs to the SSX family.
CC -!- SIMILARITY: Contains 1 KRAB-related domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X86174; CAA60110.1; -.
DR EMBL; BC001003; AAH01003.1; -.
DR EMBL; S79325; AAB35378.1; -.
DR PIR; S55057; S55057.
DR Genew; HGNC:11335; SSX1.
DR H-invDB; HIX0016770; -.
DR MIM; 312820; -.
DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0003714; F:transcription corepressor activity; TAS.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR003655; KRAB_related.
DR Pfam; PF01352; KRAB; 1.
DR SMART; SM00349; KRAB; 1.
DR PROSITE; PS50806; KRAB_RELATED; 1.
KW Chromosomal translocation; Multigene family; Proto-oncogene;
KW Transcription regulation.
FT DOMAIN 20 83 KRAB-related.
FT SITE 62 63 SSX1 fusion protein (rare).
FT SITE 110 111 SSX1-SSX1 fusion protein.
FT SITE 188 AA; 21931 MW; E440D1B2AE3AE9F7 CRC64;
SQ SEQUENCE 188 AA; 21931 MW; E440D1B2AE3AE9F7 CRC64;

Query Match 76.3%; Score 772; DB 1; Length 188;
Best Local Similarity 78.2%; Pred. No. 1.5e-53;
Matches 147; Conservative 12; Mismatches 29; Indels 0; Gaps 0;

Qy 1 MNGDDAFARRPTVGAQIPEKIQKAFDDIAKYFSKEEWMKASEKIFVYMKRYEAMTK 60
Db 1 MNGDDTFAKRRPRDDAKASEKSKAFDDIAKYFSKEEWMKMYSEKISIVYMKRYAMTK 60
```

```

QY 61 LGFKATLPPFMCNKRAEDFQGNLDNDNRGNQVERPQMTFGRLOGISPKIMPKPAEEG 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 LGFKATLPPFMCNKRAEDFQGNLDNDNRGNQVERPQMTFGRLOGISPKIMPKPAEEG 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 NDSEVPEASGPQNDGKELCPGKPTTSEKIHRSKRGHAWHRLRERKQLYVEEI 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 NDSKGVSEASGPQNDGKQLHPPGKANISEKINKRSGPKRGHAWHRLRERKQLYVEEI 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 SDPEEDDE 188
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 SDPEEDDE 188
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10
Q9BRW7 PRELIMINARY; PRT; 170 AA.
AC Q9BRW7
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Synovial sarcoma, X breakpoint 3, isoform b.
GN Name=SSX3;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=23388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Klausner R.D., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny K.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC005904; AA05904.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001909; KRAB-related.
DR Pfam; PF01352; KRAB; 1.
DR SMART; SM00349; KRAB; 1.
DR PROSITE; PS0806; KRAB RELATED; 1.
SQ SEQUENCE 170 AA; 19457 MW; D981F807A9C7EAB1 CRC64;

Query Match 72.6%; Score 735; DB 2; Length 170;
Best Local Similarity 89.1%; Pred. No. 1.1e-50;
Matches 139; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 1 MNGDDAFARRPTVGAQIPEKIQAFDDIAKYFSKEWEKMKASEKIFVYMKRKYEAMTK 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MNGDDTFARRPTVGAQIPEKIQAFDDIAKYFSKEWEKMKRVSEKIVYMKRKYEAMTK 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 11
Q96QI1 PRELIMINARY; PRT; 155 AA.
AC Q96QI1
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE DJ54B20.1.1 (Novel SSX family protein (isoform 1)) (Fragment).
GN Name=dj54B20.1;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Grafham D.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z98304; CAC41946.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001909; KRAB-related.
DR InterPro; IPR003655; KRAB-related.
DR Pfam; PF01352; KRAB; 1.
DR SMART; SM00349; KRAB; 1.
DR PROSITE; PS0806; KRAB RELATED; 1.
FT NON TER 155 155
SQ SEQUENCE 155 AA; 17689 MW; 065C7C89D35A1A30 CRC64;

Query Match 61.1%; Score 618; DB 2; Length 155;
Best Local Similarity 78.1%; Pred. No. 2e-41;
Matches 121; Conservative 10; Mismatches 24; Indels 0; Gaps 0;

QY 1 MNGDDAFARRPTVGAQIPEKIQAFDDIAKYFSKEWEKMKASEKIFVYMKRKYEAMTK 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MNGDDAFARRPTVGAQIPEKIQAFDDIAKYFSKEWEKMKFSEKISCIVHMKRKYEAMTK 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 LGFKATLPPFMCNKRAEDFQGNLDNDNRGNQVERPQMTFGRLOGISPKIMPKPAEEG 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 LGFNVTLSLFRNKRKATDSQRNDSNDNRGNQVERPQMTFGRLOGISPKIMPKPAEEG 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 NDSEVPEASGPQNDGKELCPGKPTTSEKIHRS 155
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 SDSKGVSEASGPQNDGKELCPGKASSEKIHRS 155
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12
Q96QI0 PRELIMINARY; PRT; 196 AA.
AC Q96QI0
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE DJ54B20.1.2 (Novel SSX family protein (isoform 2)) (Fragment).
GN Name=dj54B20.1;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Grafham D.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

```

```
DR EMBL; Z98304; CAC41947.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR003655; KRAB_related.
DR Pfam; PF01352; KRAB; 1.
DR SMART; SM00349; KRAB; 1.
DR PROSITE; PS0806; KRAB_RELATED; 1.
FT NON_TER 196 196
SQ SEQUENCE 196 AA; 22299 MW; 71CA1BE1F4BA07F8 CRC64;

Query Match 58.1%; Score 587.5; DB 2; Length 196;
Best Local Similarity 61.7%; Pred. No. 6.8e-39;
Matches 121; Conservative 10; Mismatches 24; Indels 41; Gaps 1;

QY 1 MNGDDAFARRPTVGAQIPEKIQK----- 23
DB 1 MNGDDAFARRPTDDAKASEKSKHPWRQVCDLALHLVLTTPFWKVGREPASITALLCGR 60
QY 24 ----AFDDIAKYFSKEWEKMKASEKIFVYVMKRYEAMTKLGFKATLPPFWCNKRAEDF 79
DB 61 GEARAFDDIAKYFSKEWEKMKFSEKISCVHMKRYEAMTKLGFNVTLSLFARNKRATDS 120
QY 80 QGNLDNDNRGNQVERPQMTFGRLOGISPKIMPKKPAEAGNDSEEVPEASGPQNDGKEL 139
DB 121 QNDSNDNRNGNEVERPQMTFGRLOGIPIPKIMPEKPAEGSDSKGVPEASGPQNDGKL 180
QY 140 CPPGKPTTSEKIHRS 155
DB 181 CPPGKASSEKIHRS 196

RESULT 13
QY Q7RTT4 PRELIMINARY; PRT; 148 AA.
AC Q7RTT4
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Synovial sarcoma X breakpoint 8 protein.
GN Name=SSX8;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22204129; PubMed=12216073;
RA Gure A.O., Wei I.J., Old L.J., Chen Y.-T.;
RT "The SSX gene family: characterization of 9 complete genes.";
RL Int. J. Cancer 101:448-453(2002).
CC -!- MISCELLANEOUS: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ third party annotation (TPA) entry.
DR EMBL; BK000688; DAA00375.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001909; KRAB.
DR InterPro; IPR003655; KRAB_related.
DR Pfam; PF01352; KRAB; 1.
DR PROSITE; PS0806; KRAB_RELATED; 1.
SQ SEQUENCE 148 AA; 17313 MW; 836B2870B89E2AA2 CRC64;

Query Match 56.0%; Score 567; DB 2; Length 148;
Best Local Similarity 77.1%; Pred. No. 2.1e-37;
Matches 111; Conservative 7; Mismatches 26; Indels 0; Gaps 0;

QY 1 MNGDDAFARRPTVGAQIPEKIQKAFDDIAKYFSKEWEKMKASEKIFVYVMKRYEAMTK 60
DB 1 MNGDDAFARRPTDDAKASEKSKAFNDIATYFSKEWEKMKYSEKISVYMKRYEAMTK 60
QY 61 LGFKATLPPFWCNKRAEDFQGNLDNDNRGNQVERPQMTFGRLOGISPKIMPKKPAEAG 120
```

```
DB 61 LGFNVTLPFPWCNKQATDFQGNVFDNRRRIQVERPQMTFGRLOGIIPKIMPKKPAEAG 120
QY 121 NDSEEVPEASGPQNDGKELCPPGK 144
DB 121 NDSKGVSEASGPQNDGKQLRRPGK 144

RESULT 14
QY Q9Y444 PRELIMINARY; PRT; 98 AA.
AC Q9Y444;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE SYT-SSX protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95038836; PubMed=7951320;
RA Clark J., Roques P.J., Crew A.J., Gill S., Shipley J., Chan A.M.,
RA Gusterson B.A., Cooper C.S.;
RT "Identification of a novel genes, SYT and SSX, involved in the t(X;18)
RT (p11.2;q11.2) translocation found in human synovial sarcoma.";
RL Nat. Genet. 7:502-508(1994).
DR EMBL; X79200; CAB36970.1; -.
FT NON_TER 1
SQ SEQUENCE 98 AA; 11136 MW; 1C71F5C8D54513DD CRC64;

Query Match 42.2%; Score 427.5; DB 2; Length 98;
Best Local Similarity 84.8%; Pred. No. 1.6e-26;
Matches 84; Conservative 1; Mismatches 13; Indels 1; Gaps 1;

QY 90 RGNQVERPQMTFGRLOGISPKIMPKKPAEAGNDSEEVPEASGPQNDGKELCPPGKPTTSE 149
DB 1 RPTQGPQPQQRPQYGYD-QIMPKKPAEAGNDSEEVPEASGPQNDGKELCPPGKPTTSE 59
QY 150 KIHRSQKRGEGHAWTHRLRERKOLVYEELSDPEEDDE 188
DB 60 KIHRSQKRGEGHAWTHRLRERKOLVYEELSDPEEDDE 98

RESULT 15
QY Q8WVZ9 PRELIMINARY; PRT; 64 AA.
AC Q8WVZ9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE BA564H18.1 (Synovial sarcoma, X breakpoint 2) (Fragment).
GN Name=SSX2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Whitehead S.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL596242; CAD18884.1; -.
FT NON_TER 1
SQ SEQUENCE 64 AA; 7014 MW; 43ADB72AF2FB9613 CRC64;

Query Match 33.1%; Score 335; DB 2; Length 64;
Best Local Similarity 96.9%; Pred. No. 2.1e-19;
Matches 62; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 95 ERPQMTFGRLOGISPKIMPKKPAEAGNDSEEVPEASGPQNDGKELCPPGKPTTSEKIH 154
```

Db 1 ERPQMTFGRLOGISPKIMPKKPAERGNDSSEVPEASGPQNDGKELCPPGKPTTSEKIHER 60

Qy 155 SGPK 158

Db 61 SGR 64

Search completed: June 20, 2005, 12:21:18  
Job time : 176 secs

**THIS PAGE LEFT BLANK**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2005, 11:55:08 ; Search time 163 Seconds  
(without alignments)  
446.079 Million cell updates/sec

Title: US-10-777-053-40  
Perfect score: 1012  
Sequence: 1 MNGDDAFARPPTVGAQIPEK.....RRKQLVIYEEISDEPDEE 188

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:.\*  
1: Geneseq1980s:.\*  
2: Geneseq1990s:.\*  
3: Geneseq2000s:.\*  
4: Geneseq2001s:.\*  
5: Geneseq2002s:.\*  
6: Geneseq2003as:.\*  
7: Geneseq2003bs:.\*  
8: Geneseq2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1012	100.0	188	5 ABG95088	Abg95088 Human tra
2	1012	100.0	188	5 AAU1542	Aau1542 Human Mel
3	1012	100.0	188	6 ABP74130	Abp74130 Human SSX
4	1012	100.0	188	6 ABR83437	AbR83437 Human SSX
5	1012	100.0	188	7 ADC09567	Adc09567 SSX-2 pro
6	1012	100.0	188	7 ADD25524	Add25524 Binding d
7	1012	100.0	188	8 ADJ54141	Adj54141 Human SSX
8	1012	100.0	188	8 ADM72744	Adm72744 Human SSX
9	1012	100.0	188	8 ADQ10475	Adq10475 Human tum
10	1000	98.8	188	8 ADK68677	Adk68677 Epitope l
11	998	97.6	188	2 AAR90677	Aar90677 Human X-C
12	911	90.0	207	8 ADQ10464	Adq10464 Epitope l
13	910.5	90.0	208	8 ADQ10463	Adq10463 Epitope l
14	910	89.9	169	8 ADQ10460	Adq10460 Human tum
15	910	89.9	245	8 ADQ10461	Adq10461 Epitope l
16	910	89.9	245	8 ADQ10462	Adq10462 Epitope l
17	903	89.2	245	8 ADK68664	Adk68664 Epitope l
18	902	89.1	207	8 ADK68666	Adk68666 Epitope l
19	900.5	89.0	208	8 ADK68665	Adk68665 Epitope l
20	900	88.9	245	8 ADK68663	Adk68663 Epitope l
21	894.5	88.4	170	8 ADK68662	Adk68662 Epitope l
22	837	82.7	175	4 ABG16215	Abg16215 Novel hum
23	837	82.7	175	4 ABG16212	Abg16212 Novel hum
24	814	80.4	305	8 ADN99849	Adn99849 Novel hum
25	798	78.9	188	6 ABP74710	Abp74710 Human SSX

26	798	78.9	188	7 ADC09597	Adc09597 SSX-4 pro
27	798	78.9	188	8 ADJ54142	Adj54142 Human SSX
28	798	78.9	188	8 ADM72835	Adm72835 Human SSX
29	798	78.9	188	8 ADL14098	Adl14098 Human sar
30	772	76.3	188	2 AAR90676	Aar90676 Human X-C
31	772	76.3	188	8 ADL14092	Adl14092 Human sar
32	427.5	42.2	98	2 AAR90678	Aar90678 Human SYT
33	363.5	35.9	492	7 ADJ69335	Adj69335 Human hea
34	343.3	34.9	366	8 ADN99661	Adn99661 Novel hum
35	349.5	34.5	158	2 AAR90679	Aar90679 Human SYT
36	294	29.1	143	4 ABG16217	Abg16217 Novel hum
37	291	28.8	145	4 ABG16214	Abg16214 Novel hum
38	265	26.2	55	5 ABG95087	Abg95087 Human tra
39	201	19.9	38	6 ABP74134	Abp74134 Human SSX
40	201	19.9	38	7 ADC09003	Adc09003 Epitope w
41	201	19.9	38	8 ADM72751	Adm72751 Human SSX
42	185	18.3	35	8 ADM73366	Adm73366 Human SSX
43	176	17.4	33	8 ADM73365	Adm73365 Human SSX
44	164	16.2	31	8 ADM73361	Adm73361 Human SSX
45	151	14.9	369	4 ABG27791	Abg27791 Novel hum

## ALIGNMENTS

## RESULT 1

ABG95088  
ID ABG95088 standard; protein; 188 AA.

XX AC ABG95088;

DT 04-DEC-2002 (first entry)

DE Human translocation (X; 18) (p11.2; q11.2) protein #2.

XX Chromosome aberration; oncogenic fusion protein; cancer;

KW proliferative disease; cellular protein isoform; heat shock protein 90;

KW HSP-90; rheumatoid arthritis; cancer; haematopoietic disorder;

KW T cell lymphoma; B cell lymphoma; chronic myeloid leukaemia; CML;

KW acute myeloid leukaemia; AML; chronic myelomonocytic leukaemia; CMML;

KW lymphoblastic leukaemia; ALL; APL; NHL; solid tumour;

KW papillary thyroid carcinoma; Ewing's sarcoma; melanoma; liposarcoma;

KW rhabdomyosarcoma; synovial sarcoma; viral infection.

XX Homo sapiens.

XX WO200269900-A2.

PD 12-SEP-2002.

PF 01-MAR-2002; 2002WO-US006518.

PR 01-MAR-2001; 2001US-0272751P.

XX (CONF-) CONFORMA THERAPEUTICS CORP.

XX Fritz LC, Burrows FJ;

XX WPI; 2002-698710/75.

XX N-PSDB; ABS73281.

XX Treating genetically-defined disease associated with chromosomal aberrations yielding oncogenic fusion proteins, e.g. cell proliferative diseases, involves administering an inhibitor of heat shock protein 90.

XX Disclosure; Page 236; 389pp; English.

XX The invention describes a method of treating genetically-defined disease associated with chromosomal aberrations yielding oncogenic fusion proteins (I), treating cancerous cells containing (I) in a heterogeneous cell population, treating proliferative diseases associated with mutant protein or cellular protein isoforms (II) dependent on heat shock protein (HSP)-90, or selectively treating cells expressing (II) involving

CC administering HSP90-inhibitor. The method is useful for treating  
CC genetically-defined disease with chromosomal aberration yielding  
CC oncogenic fusion protein, treating cancerous cells containing fusion  
CC protein in heterogeneous cell population, treating proliferative disease  
CC (e.g. rheumatoid arthritis or cancer) associated with mutant protein or  
CC cellular protein isoform dependent on heat shock protein (HSP)-90 (e.g.  
CC p53), or selectively treating cells expressing mutant protein or cellular  
CC protein isoform in a patient heterozygous for (II). The method is useful  
CC for treating a disease e.g. haematopoietic disorder such as T or B cell  
CC lymphoma, chronic myeloid leukaemia (CML), APL, ALL, NHL and CML,  
CC or a disease characterised by a solid tumour such as papillary thyroid  
CC carcinoma, Ewing's sarcoma, melanoma, liposarcoma, rhabdomyosarcoma and  
CC synovial sarcoma. The method is also useful for treating viral  
CC infections. This represents a protein encoded by the DNA sequence of a  
CC chromosome aberration  
XX  
SQ Sequence 188 AA;  
Query Match 100.0%; Score 1012; DB 5; Length 188;  
Best Local Similarity 100.0%; Pred. No. 7e-91;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MNGDDAFARRPTVGAQIPEKIQAFDDIAKYFSEWEKMKASEKIFYVMKRYEAMTK 60  
DB 1 MNGDDAFARRPTVGAQIPEKIQAFDDIAKYFSEWEKMKASEKIFYVMKRYEAMTK 60  
QY 61 LGFKATLPPFCMKRAEDFQGNLDNDPNRGQVVERPQMTFGRLOGISPKIMPKPAEAG 120  
DB 61 LGFKATLPPFCMKRAEDFQGNLDNDPNRGQVVERPQMTFGRLOGISPKIMPKPAEAG 120  
QY 121 NDSEEVPEASGPNQDQKELCPGKPTTSKIHRSQPKRGEHAWTHRLRERKQLVYEEI 180  
DB 121 NDSEEVPEASGPNQDQKELCPGKPTTSKIHRSQPKRGEHAWTHRLRERKQLVYEEI 180  
QY 181 SDPEEDDE 188  
DB 181 SDPEEDDE 188  
RESULT 2  
AAU11542  
ID AAU11542 standard; protein; 188 AA.  
XX  
AC AAU11542;  
DT 12-MAR-2002 (first entry)  
XX  
DE Human Melanoma tumour associated antigen SSX-2.  
XX  
KW Human; melanoma tumour associated antigen; human leukocyte antigen;  
KW major histocompatibility complex; MHC; HLA-A2.2; vaccine; cancer; HIV;  
KW human immunodeficiency virus infection; cytostatic; virucide;  
KW housekeeping epitope; adoptive immunotherapy; neoplastic disease;  
KW viral disease; hepatitis virus; papilloma virus; tumour; leukaemia;  
KW lymphoma; breast cancer; prostate cancer; lung cancer; SSX-2;  
KW parasitic infection; Chlamydia; Trypanosoma; Toxoplasma.  
XX  
OS Homo sapiens.  
XX  
FN WO200182963-A2.  
PD 08-NOV-2001.  
PF 27-APR-2001; 2001WO-US013806.  
XX  
PR 28-APR-2000; 2000US-00560465.  
PR 28-APR-2000; 2000US-00561074.  
PR 28-APR-2000; 2000US-00561571.  
PR 28-APR-2000; 2000US-00561572.  
XX  
PA (CTLI-) CTL IMMUNOTHERAPIES CORP.  
XX  
PI Simard JJJ, Diamond DC, Lei X;

XX WPI; 2002-066492/09.  
DR Novel vaccine useful for treating neoplastic and viral diseases,  
XX comprises a first housekeeping epitope derived from a first antigen  
PT associated with a first target cell.  
PT  
XX Example 22; Fig 16; 131pp; English.  
XX  
CC The invention relates to a vaccine comprising a first housekeeping  
CC epitope derived from a first antigen associated with a first target cell.  
CC Also included are an isolated T cell expressing a T cell receptor  
CC specific for a major histocompatibility complex (MHC)-peptide complex  
CC comprising a first housekeeping epitope which is derived from a first  
CC antigen associated with a first target cell, selecting an epitope (or  
CC peptide sequence) from a population of peptide fragments of an antigen  
CC associated with a target in a host, where the fragments have a known or  
CC predicted affinity for a MHC receptor peptide binding cleft of the host,  
CC where the epitope selected corresponds to a product of proteolytic  
CC cleavage of the antigen in a cell of the host and a nucleic acid  
CC construct comprising a first coding region, where the first coding region  
CC comprises a first sequence encoding at least a first polypeptide, where  
CC the first polypeptide comprises a first housekeeping epitope derived from  
CC a first antigen associated with a first target cell; the epitopes,  
CC peptides, vaccines and nucleic acids are useful in the manufacture of a  
CC medicament for use in adoptive immunotherapy and for prevention and  
CC treatment of neoplastic and viral diseases (e.g. human immunodeficiency  
CC virus, HIV, infection, hepatitis virus and papilloma virus), cancers  
CC (e.g. tumours, leukaemia, lymphoma, breast cancer, prostate cancer and  
CC lung cancer), infection of cells by intracellular parasites (e.g.  
CC Chlamydia, Trypanosoma and Toxoplasma) and many other examples given in  
CC the specification. The invention permits the vaccine designer to ignore  
CC peptides that, despite predicted high binding affinity for MHC, will  
CC never be useful because they cannot be presented by target cells. The  
CC invention provides a major advance in vaccine design, one that combines  
CC the power of antigen sequence analysis with the fundamental realities of  
CC immunology. The invention allows for the simple and effective selection  
CC of meaningful epitopes for creation of MHC class I or Class II vaccines  
CC using any polypeptide sequence corresponding to a desired target. The  
CC present sequence is an HLA-A2.1 (human leukocyte antigen) presenting  
CC target cell protein from which epitopes of the invention may be derived,  
CC SSX-2 (a melanoma tumour associated antigen)  
XX  
SQ Sequence 188 AA;  
Query Match 100.0%; Score 1012; DB 5; Length 188;  
Best Local Similarity 100.0%; Pred. No. 7e-91;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MNGDDAFARRPTVGAQIPEKIQAFDDIAKYFSEWEKMKASEKIFYVMKRYEAMTK 60  
DB 1 MNGDDAFARRPTVGAQIPEKIQAFDDIAKYFSEWEKMKASEKIFYVMKRYEAMTK 60  
QY 61 LGFKATLPPFCMKRAEDFQGNLDNDPNRGQVVERPQMTFGRLOGISPKIMPKPAEAG 120  
DB 61 LGFKATLPPFCMKRAEDFQGNLDNDPNRGQVVERPQMTFGRLOGISPKIMPKPAEAG 120  
QY 121 NDSEEVPEASGPNQDQKELCPGKPTTSKIHRSQPKRGEHAWTHRLRERKQLVYEEI 180  
DB 121 NDSEEVPEASGPNQDQKELCPGKPTTSKIHRSQPKRGEHAWTHRLRERKQLVYEEI 180  
QY 181 SDPEEDDE 188  
DB 181 SDPEEDDE 188  
RESULT 3  
ABP74130  
ID ABP74130 standard; protein; 188 AA.  
XX  
AC ABP74130;  
XX  
DT 03-FEB-2003 (first entry)

```
XX DE Human SX-2 protein SEQ ID NO:3.
XX AC
XX DT 13-OCT-2003 (first entry)
XX DE Human SX-2 amino acid sequence SEQ ID NO:2.
XX OS Homo sapiens.
XX PN WO2002081646-A2.
XX KW Epitope cluster; cytostatic; virucide; antibacterial; protozoacide;
XX KW fungicide; vaccine; adoptive immunotherapy; immune response; cancer;
XX KW intracellular parasite; infection; immunotherapeutic; neoplastic disease;
XX KW viral disease; human; carcinoembryonic antigen; CEA.
XX OS
XX PN WO2003057823-A2.
XX PD 17-JUL-2003.
XX PF 05-NOV-2002; 2002WO-US035582.
XX PR 07-NOV-2001; 2001US-00005905.
XX PR 07-DEC-2001; 2001US-00026066.
XX PA (CTLI-) CTL IMMUNOTHERAPIES CORP.
XX PI Simard JLL, Diamond DC;
XX WPI; 2003-577518/54.
XX PT New epitope cluster derived from an antigen associated with a target.
XX PT useful for treating or preventing neoplastic or viral diseases, or
XX PT inducing an immune response against cancer cells or cells infected with
XX PT intracellular parasites.
XX PS Example 22; Fig 16; 190pp; English.
XX CC The present invention describes an epitope cluster (I) derived from an
XX CC antigen associated with a target, where the cluster is a fragment of the
XX CC antigen. The cluster comprises or encodes at least two sequences having a
XX CC known or predicted affinity for a major histocompatibility complex
XX CC receptor-binding cleft. (I) has the structure: X-P21-XAN-P2N-X(bN-1)-P1-
XX CC XAN-PN, where X = any amino acid naturally occurring in a protein
XX CC sequence; Xa and X(b/-1) = strings of amino acids of length a and b/-1,
XX CC respectively; a = the number of amino acids between P21 and P2N; (b-1) =
XX CC the number of amino acids between P2N and P1; P21 = a first primary
XX CC anchor and second residue of a first epitope; P2N = a first primary
XX CC anchor and second residue of an Nth epitope; P1 = a last primary anchor
XX CC and C-terminal residue of the first epitope; 2N, Nc, N = the Nth epitope
XX CC of the cluster; Nc = the total number of epitopes in the cluster; and AN
XX CC and bN = the positional relationship between the first and Nth epitope.
XX CC (I) has cytostatic, virucide, antibacterial, protozoacide and fungicide,
XX CC and can be used in vaccines. T cells from the present invention can be
XX CC used for manufacturing a medicament for adoptive immunotherapy. The
XX CC epitope clusters, polypeptides, polynucleotides or vaccines from the
XX CC present invention can be used for inducing an immune response against
XX CC cancer cells or cells infected with intracellular parasites such as
XX CC viruses, bacteria, protozoa, fungi or prion. The T cell or
XX CC immunotherapeutic composition from the present invention can be used for
XX CC treating or preventing neoplastic or viral diseases. The present sequence
XX CC is used in the exemplification of the present invention
XX SQ Sequence 188 AA;
XX Query Match 100.0%; Score 1012; DB 6; Length 188;
XX Best Local Similarity 100.0%; Pred. No. 7e-91;
XX Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 MNGDDAFARRPTVGAQIPEKIOKAFDDIAKYFSKEWEKAKASEKIFYVMKKRYEAMTK 60
XX DB 1 MNGDDAFARRPTVGAQIPEKIOKAFDDIAKYFSKEWEKAKASEKIFYVMKKRYEAMTK 60
XX QY 61 LGFKATLPPFMCNKRAEDFGNDLNDNRGNQVERPQMTFGRLQGISPKIMPKPAERG 120
XX DB 61 LGFKATLPPFMCNKRAEDFGNDLNDNRGNQVERPQMTFGRLQGISPKIMPKPAERG 120
XX QY 121 NSEEVPEASGPNQDKELCPGKPTTSEKIHRSQPKGEHAWTHRLRERKQLVYEI 180
XX DB 121 NSEEVPEASGPNQDKELCPGKPTTSEKIHRSQPKGEHAWTHRLRERKQLVYEI 180
XX QY 181 SDPEEDE 188
XX DB 181 SDPEEDE 188
XX RESULT 4
XX ABR83437
XX ID ABR83437 standard; protein; 188 AA.
```

Db 61 LGFKATLPPFMCNKRAEDFQGNLDNDPNRGNQVERPQMTFGRLOGISPKIMPKPAEAG 120  
QY 121 NDSEEVPEASGPQNDGKELCPGKPTTSEKIHRSKPGEHAWTHRLRERKQLVIYEEI 180  
Db 121 NDSEEVPEASGPQNDGKELCPGKPTTSEKIHRSKPGEHAWTHRLRERKQLVIYEEI 180  
QY 181 SDPEEDE 188  
Db 181 SDPEEDE 188

RESULT 5  
ADC09567  
ID ADC09567 standard; protein; 188 AA.  
XX  
AC ADC09567;  
DT 18-DEC-2003 (first entry)  
XX  
DE SSX-2 protein #SEQ ID 3.  
XX  
KW Epitope; immunological; vaccine;  
KW major histocompatibility complex class I; MHC class I; cancer;  
KW immunisation.  
XX  
OS Unidentified.  
XX  
PN WO2003008537-A2.  
XX  
PD 30-JAN-2003.  
XX  
PF 29-MAR-2002; 2002WO-US010189.  
XX  
PR 06-APR-2001; 2001US-0282211P.  
PR 07-NOV-2001; 2001US-0337017P.  
PR 07-MAR-2002; 2002US-0363210P.  
XX  
PA (CTL1-) CTL IMMUNOTHERAPIES CORP.  
XX  
PI Simard JLL, Diamond DC, Liu L, Xie Z;  
XX  
XX WPI; 2003-248010/24.  
DR  
XX  
XX Epitope having high affinity for major histocompatibility complex class I  
PT useful for treating an animal, evaluating immunogenicity of a vaccine or  
PT therapeutic composition and for diagnosing a disease.  
XX  
PS Claim 1; SEQ ID NO 3; 239pp; English.  
XX  
XX The invention relates to an isolated epitope polypeptide that has high  
CC affinity for major histocompatibility complex (MHC) class I, and an  
CC epitope cluster comprising the polypeptide. Also disclosed is a vaccine  
CC or immunotherapeutic composition containing an epitope of the invention.  
CC Compositions of the invention may be used in the treatment of cancer. The  
CC compositions of the invention may be used in the treatment of cancer. The  
CC method can be combined with a radiation therapy, chemotherapy,  
CC biochemotherapy or surgery. The composition is also useful for evaluating  
CC immunogenicity of a vaccine or immunotherapeutic compound. Multimeric MHC  
CC peptide complexes of the invention are useful for determining specific T  
CC cell frequency. This method is useful for evaluating immunological  
CC response, by performing the method prior to and subsequent to an  
CC immunisation step. Compositions of the invention are useful for  
CC diagnosing a disease. The current sequence represents an epitope of the  
CC invention with high affinity for MHC class I.  
XX  
SQ Sequence 188 AA;

Query Match 100.0%; Score 1012; DB 7; Length 188;  
Best Local Similarity 100.0%; Pred. No. 7e-91;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNGDDAFARRPTVGAQIPEKIQAFFDDIAKYSKEWEKKASEKIFYYVMKKRYEAMTK 60  
Db 1 MNGDDAFARRPTVGAQIPEKIQAFFDDIAKYSKEWEKKASEKIFYYVMKKRYEAMTK 60

QY 61 LGFKATLPPFMCNKRAEDFQGNLDNDPNRGNQVERPQMTFGRLOGISPKIMPKPAEAG 120  
Db 61 LGFKATLPPFMCNKRAEDFQGNLDNDPNRGNQVERPQMTFGRLOGISPKIMPKPAEAG 120  
QY 121 NDSEEVPEASGPQNDGKELCPGKPTTSEKIHRSKPGEHAWTHRLRERKQLVIYEEI 180  
Db 121 NDSEEVPEASGPQNDGKELCPGKPTTSEKIHRSKPGEHAWTHRLRERKQLVIYEEI 180  
QY 181 SDPEEDE 188  
Db 181 SDPEEDE 188

RESULT 6  
ADD25524  
ID ADD25524 standard; protein; 188 AA.  
XX  
AC ADD25524;  
DT 15-JAN-2004 (first entry)  
XX  
DE Binding domain-immunoglobulin fusion protein-associated protein #39.  
XX  
KW Binding domain; immunoglobulin; fusion protein; cytostatic;  
KW antiarthritic; immunosuppressive; antidiabetic; antithyroid;  
KW neuroprotective; hinge region; immunoglobulin heavy chain;  
KW CH2 constant region; CH3 constant region; IgG1;  
KW antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation;  
KW malignant condition; B-cell disorder; melanoma; carcinoma; sarcoma;  
KW rheumatoid arthritis; myasthenia gravis; Grave's disease;  
KW type I diabetes mellitus; multiple sclerosis; autoimmune disease.  
XX  
OS Unidentified.  
XX  
PN US2003118592-A1.  
XX  
PD 26-JUN-2003.  
XX  
PF 25-JUL-2002; 2002US-00207655.  
XX  
PR 17-JAN-2001; 2001US-0367358P.  
PR 17-JUN-2002; 2002US-00053530.  
PR 03-JUN-2002; 2002US-0385691P.  
XX  
XX (GENE-) GENE-CRAFT INC.  
XX  
PI Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;  
XX  
XX WPI; 2003-801317/75.  
DR  
XX  
XX New binding domain-immunoglobulin fusion protein, useful for treating a  
PT subject having or suspected of having a malignant condition or a B-cell  
PT disorder, e.g. melanoma, Grave's disease or autoimmune disease.  
XX  
PS Disclosure; SEQ ID NO 85; 157pp; English.  
XX  
XX The invention relates to a binding domain-immunoglobulin fusion protein  
CC comprising a binding domain polypeptide that is fused to an  
CC immunoglobulin hinge region polypeptide, an immunoglobulin heavy chain  
CC CH2 constant region polypeptide that is fused to the hinge region  
CC polypeptide, and an immunoglobulin heavy chain CH3 constant region  
CC polypeptide that is fused to the CH2 constant region polypeptide. The  
CC hinge region polypeptide comprises: a wild-type human IgG1 immunoglobulin  
CC hinge region polypeptide; a mutated human IgG1 immunoglobulin hinge  
CC region polypeptide, derived from (a) having 3 or more cysteine residues;  
CC where the mutated human IgG1 immunoglobulin hinge region polypeptide  
CC contains 2 cysteine residues, where the first cysteine is not mutated; a  
CC mutated human IgG1 immunoglobulin hinge region polypeptide, derived from  
CC (a) having 3 or more cysteine residues, where the mutated human IgG1  
CC immunoglobulin hinge region polypeptide contains no more than one  
CC cysteine residue; and a mutated human IgG1 immunoglobulin hinge region  
CC polypeptide, derived from (a) having 3 or more cysteine residues; where

CC the mutated human IgG1 immunoglobulin hinge region polypeptide contains  
CC no cysteine residues. The binding domain-immunoglobulin fusion protein is  
CC capable of at least one immunological activity comprising antibody  
CC dependent cell-mediated cytotoxicity (ADCC) and complement fixation. The  
CC binding domain polypeptide is capable of specifically binding to an  
CC antigen. Also included are an isolated polynucleotide encoding the  
CC binding domain-immunoglobulin fusion protein, a recombinant expression  
CC construct comprising the polynucleotide (operably linked to a promoter),  
CC a host cell transformed or transfected with a recombinant expression  
CC construct, producing the binding domain-immunoglobulin fusion protein, a  
CC pharmaceutical composition comprising the binding domain-immunoglobulin  
CC fusion protein or polynucleotide and a carrier, and treating a subject  
CC having or suspected of having a malignant condition or a B-cell disorder.  
CC The binding domain-immunoglobulin fusion protein is useful for treating a  
CC subject having or suspected of having a malignant condition or a B-cell  
CC disorder, e.g. melanoma, carcinoma or sarcoma, rheumatoid arthritis,  
CC myasthenia gravis, Grave's disease, type I diabetes mellitus, multiple  
CC sclerosis or autoimmune disease. The present sequence is a binding domain  
CC -immunoglobulin fusion protein-associated protein sequence. Note: The  
CC sequence data for this patent formed part of the printed specification  
CC and is also available in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?DocId=20030118592. The authors have not  
CC identified the sequences in the printed specification by their SEQ ID  
CC number therefore none of the sequences can be explicitly identified.  
XX  
XX

XX Sequence 188 AA;

Query Match 100.0%; Score 1012; DB 7; Length 188;  
Best Local Similarity 100.0%; Pred. No. 7e-91;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MNGDDAFARRPTVGAQIPEKIOKAFDDIAKYFSKEWEKWKASEKIFYVMKRYEAMTK 60  
DB 1 MNGDDAFARRPTVGAQIPEKIOKAFDDIAKYFSKEWEKWKASEKIFYVMKRYEAMTK 60  
QY 61 LGFKATLPPFMCNKRAEDFQGNLDNDPNRGNQVERPQMTFGRLOGISPKIMPKPAERG 120  
DB 61 LGFKATLPPFMCNKRAEDFQGNLDNDPNRGNQVERPQMTFGRLOGISPKIMPKPAERG 120  
QY 121 NDSEVPEASGPQNDGKELCPGKPTTSKIHRSKPKRGEHAWTHRLRERKOLVYEEI 180  
DB 121 NDSEVPEASGPQNDGKELCPGKPTTSKIHRSKPKRGEHAWTHRLRERKOLVYEEI 180  
QY 181 SDPEEDDE 188  
DB 181 SDPEEDDE 188

RESULT 7  
ADJ54141  
ID ADJ54141 standard; protein; 188 AA.  
XX AC ADJ54141;  
DT 06-MAY-2004 (first entry)  
XX DE Human SSX2 polypeptide.  
XX KW Human; lung cancer; bromo domain testis specific gene product; BRDT;  
KW seminoma; melanoma; teratoma; glioma; ovarian cancer; colorectal cancer;  
KW SSX2.  
XX OS Homo sapiens.  
XX PN US6686147-B1.  
XX PD 03-FEB-2004.  
XX PF 09-SEP-1999; 99US-00392714.  
XX PR 15-JUL-1998; 98WO-US014679.  
XX PA (LUDW-) LUDWIG INST CANCER RES.

XX Scanlan MJ, Gure A, Old LJ, Chen Y, Williamson B;  
XX WPI; 2004-223796/21.  
DR N-PSDB; ADJ54130.  
XX Diagnosing lung cancer by contacting sample comprising lung cells from  
PT subject with probe that hybridizes to nucleic acid having human bromo  
PT domain testis specific gene product cDNA sequence and determining  
PT hybridization.  
XX Disclosure; SEQ ID NO 27; 59pp; English.  
PS The invention relates to a method of diagnosing lung cancer by contacting  
XX a biological sample comprising lung cells from a subject, with a probe  
CC that hybridizes under high-stringency conditions to a nucleic acid  
CC molecule having a fully defined human bromo domain testis (BRDT) specific  
CC gene product cDNA sequence and determining expression of the nucleic acid  
CC molecule or its complement in sample. The method is useful for the  
CC preparation of a medicament for treating cancer, preferably seminomas,  
CC melanomas, teratomas, gliomas, lung cancer, ovarian cancer and/or  
CC colorectal cancer. This sequence represents the human SSX2 polypeptide,  
CC used in the scope of the invention.  
XX Sequence 188 AA;

Query Match 100.0%; Score 1012; DB 8; Length 188;  
Best Local Similarity 100.0%; Pred. No. 7e-91;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNGDDAFARRPTVGAQIPEKIOKAFDDIAKYFSKEWEKWKASEKIFYVMKRYEAMTK 60  
DB 1 MNGDDAFARRPTVGAQIPEKIOKAFDDIAKYFSKEWEKWKASEKIFYVMKRYEAMTK 60  
QY 61 LGFKATLPPFMCNKRAEDFQGNLDNDPNRGNQVERPQMTFGRLOGISPKIMPKPAERG 120  
DB 61 LGFKATLPPFMCNKRAEDFQGNLDNDPNRGNQVERPQMTFGRLOGISPKIMPKPAERG 120  
QY 121 NDSEVPEASGPQNDGKELCPGKPTTSKIHRSKPKRGEHAWTHRLRERKOLVYEEI 180  
DB 121 NDSEVPEASGPQNDGKELCPGKPTTSKIHRSKPKRGEHAWTHRLRERKOLVYEEI 180  
QY 181 SDPEEDDE 188  
DB 181 SDPEEDDE 188

RESULT 8  
ADM72744  
ID ADM72744 standard; protein; 188 AA.  
XX AC ADM72744;  
DT 03-JUN-2004 (first entry)  
XX DE Human SSX-2 protein SEQ ID NO:3.  
XX KW epitope; epitope cluster; virucide; cytostatic; vaccine; viral infection;  
KW cancer; tumour; human; SSX-2.  
XX OS Homo sapiens.  
XX PN WO2004022709-A2.  
XX PD 18-MAR-2004.  
XX PF 05-SEP-2003; 2003WO-US027706.  
XX PR 06-SEP-2002; 2002US-0409123P.  
XX PA (MANN-) MANNKIND CORP.  
XX PI Simard JJJ, Diamond DC, Liu L, Liu Z;

```

XX WPI; 2004-315564/29.
DR N-P9DB; ADM72747.
XX
XX
PT New polypeptides and encoding nucleic acids that are useful epitopes of
PT target-associated antigens, useful for diagnosing and/or treating viral
PT infections, cancers and tumors.
XX
XX Disclosure; SEQ ID NO 3; 357pp; English.
PS
XX
CC The present invention describes a polypeptide (I) comprising a component
CC selected from: (a) a polypeptide epitope having any of the 503 fully
CC defined sequences of 8-33 amino acids (SEQ ID NO:108-610); (b) an epitope
CC cluster comprising the polypeptide of (a); (c) a polypeptide having
CC substantial similarity to (a) or (b); (d) a polypeptide having functional
CC similarity to any of (a)-(c); or (e) a nucleic acid encoding the
CC polypeptide of (a)-(c). (I) has virucide and cytostatic activities, and
CC can be used in vaccines. The methods and compositions of the present
CC invention are useful for the diagnosis and/or treatment of viral
CC infections, cancers and tumours. The present sequence is used in the
CC exemplification of the present invention.
XX
XX Sequence 188 AA;
SQ
    Query Match      100.0%; Score 1012; DB 8; Length 188;
    Best Local Similarity 100.0%; Pred. No. 7e-91;
    Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    .QY 1 MNGDDAFARRPTVGAQIPEKIQKAFDDIAKYFSKEWEKOKASEKIFYYMKRKYEAMTK 60
    Db 1 MNGDDAFARRPTVGAQIPEKIQKAFDDIAKYFSKEWEKOKASEKIFYYMKRKYEAMTK 60
    QY 61 LGFKATLPPFMCNKRAEDFQGNLDNDPNRGQVERPQMTFGRLOGISPKIMPKPAEAG 120
    Db 61 LGFKATLPPFMCNKRAEDFQGNLDNDPNRGQVERPQMTFGRLOGISPKIMPKPAEAG 120
    QY 121 NDSEEVPEASGPQNDGKELCPGKPTTSKIHRSKPGKGEHAWTHRLRERKQLVYEEI 180
    Db 121 NDSEEVPEASGPQNDGKELCPGKPTTSKIHRSKPGKGEHAWTHRLRERKQLVYEEI 180
    QY 181 SDPEEDE 188
    Db 181 SDPEEDE 188
    RESULT 9
    ADQ10475
    ID ADQ10475 standard; protein; 188 AA.
    XX ADQ10475;
    XX
    DT 23-SEP-2004 (first entry)
    XX
    DE Human tumour-associated antigen SSX-2.
    XX
    KW immunostimulant; cytostatic; vaccine; tumour-associated antigen SSX-2;
    KW SSX-2 antigen; epitope cluster; MHC receptor peptide binding cleft;
    KW immunogenic composition; immune response; cancer; vaccine vector;
    KW epitope liberation; human leukocyte antigen; HLA A2-specific CTL;
    KW . cytotoxic T lymphocyte; human; synovial sarcoma X breakpoint 2.
    XX
    OS Homo sapiens.
    XX
    PN US2004132088-A1.
    XX
    PD 08-JUL-2004.
    XX
    PF 10-FEB-2004; 2004US-00777053.
    XX
    PR 07-NOV-2001; 2001US-0336968P.
    PR 07-NOV-2002; 2002US-00292413.
    XX
    ,PA (SIMA/) SIMARD J J L.
    PA
    (DIAM/) DIAMOND D C.
    PA (QIUZ/) QIU Z.
    PA (LEIX/) LEI X.
    XX
    PI Simard J J L, Diamond DC, Qiu Z, Lei X;
    XX
    DR WPI; 2004-517003/49.
    DR N-PSDB; ADQ10476.
    XX
    PT Novel nucleic acid encoding tumor-associated antigen SSX-2, useful in
    PT inducing an immune response and in treating cancer.
    PS
    PS Claim 1; SEQ ID NO 40; 260pp; English.
    CC
    CC The invention describes an isolated nucleic acid (I) comprising a reading
    CC frame comprising a first sequence, where the first sequence encodes one
    CC or more segments of tumour-associated antigen SSX-2, which comprises a
    CC sequence of 188 amino acids (SEQ ID NO: 40), where the first sequence
    CC does not encode the complete SSX-2 antigen, and where each segment
    CC comprises an epitope cluster, the cluster comprising or encoding at least
    CC two amino acid sequences having a known or predicted affinity for a same
    CC MHC receptor peptide binding cleft. Also described are: an isolated
    CC polypeptide comprising the amino acid sequence encoded in the reading
    CC frame; and an immunogenic composition comprising (I) or the polypeptide
    CC of (I). (I) is a nucleic acid encoding a tumour-associated antigen SSX-2
    CC comprising a fully defined sequence of 188 amino acids (SEQ ID NO: 40).
    CC The nucleic acid, the encoded antigen, and composition are useful in
    CC inducing an immune response and in treating cancer. Expression cassettes
    CC are used in vaccine vectors. This is the amino acid sequence of human
    CC tumour-associated antigen SSX-2.
    XX
    XX Sequence 188 AA;
    Query Match      100.0%; Score 1012; DB 8; Length 188;
    Best Local Similarity 100.0%; Pred. No. 7e-91;
    Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    .QY 1 MNGDDAFARRPTVGAQIPEKIQKAFDDIAKYFSKEWEKOKASEKIFYYMKRKYEAMTK 60
    Db 1 MNGDDAFARRPTVGAQIPEKIQKAFDDIAKYFSKEWEKOKASEKIFYYMKRKYEAMTK 60
    QY 61 LGFKATLPPFMCNKRAEDFQGNLDNDPNRGQVERPQMTFGRLOGISPKIMPKPAEAG 120
    Db 61 LGFKATLPPFMCNKRAEDFQGNLDNDPNRGQVERPQMTFGRLOGISPKIMPKPAEAG 120
    QY 121 NDSEEVPEASGPQNDGKELCPGKPTTSKIHRSKPGKGEHAWTHRLRERKQLVYEEI 180
    Db 121 NDSEEVPEASGPQNDGKELCPGKPTTSKIHRSKPGKGEHAWTHRLRERKQLVYEEI 180
    QY 181 SDPEEDE 188
    Db 181 SDPEEDE 188
    RESULT 10
    ADK68677
    ID ADK68677 standard; protein; 188 AA.
    XX
    AC ADK68677;
    XX
    DT 06-MAY-2004 (first entry)
    XX
    DE Epitope liberation-related SSX-2 protein SeqID40.
    XX
    KW epitope liberation; substrate; proteasome; cytostatic; antibacterial;
    KW protozoacide; fungicide; T-cell activator; vaccine; housekeeping epitope;
    KW cytotoxic T lymphocyte; CTL; adoptive immunotherapy; neoplastic cell;
    KW virus; bacterium; protozoan; fungus; housekeeping proteasome system;
    KW human.
    XX
    OS Homo sapiens.
    XX
    PN US2003228634-A1.

```

XX 11-DEC-2003.  
XX 07-NOV-2002; 2002US-00292413.  
XX 07-NOV-2001; 2001US-0336968P.  
XX (SIMA/J) SIMARD J J L.  
XX (DIAM/J) DIAMOND D C.  
XX (QIUZ/J) QIU Z.  
XX (LEIX/J) LEI X.  
XX Simard J J L, Diamond DC, Qiu Z, Lei X;  
XX WPI: 2004-167209/16.  
XX N-PSDB; ADK68678.  
XX Identifying polypeptide suitable for epitope e.g., housekeeping epitope,  
XX liberation by contacting substrate polypeptide comprising epitope of  
XX interest, with proteasome, and assaying for liberation of epitope.  
XX Disclosure; SEQ ID NO 40; 67pp; English.  
XX This invention relates to a novel method of identifying a polypeptide  
XX suitable for epitope liberation, including the steps of identifying an  
XX epitope of interest; providing substrate polypeptide sequence including  
XX the epitope, wherein the substrate permits processing by a proteasome;  
XX contacting the substrate with a composition including the proteasome,  
XX under conditions that support processing of the substrate by proteasome;  
XX and assaying for liberation of epitope. The invention may be useful for  
XX the development of compounds with a cytostatic, antibacterial,  
XX protozoacide or fungicide activity acting as T-cell activators. In  
XX addition, the invention may allow development of a vaccine. The invention  
XX is useful for identifying a polypeptide suitable for epitope liberation,  
XX where the epitope is a housekeeping epitope. The compositions comprising  
XX the identified housekeeping epitopes are useful in vitro in vaccine  
XX development or in the generation or expansion of cytotoxic T lymphocyte  
XX (CTL) to be used in adoptive immunotherapy. The invention is also useful  
XX for activating T-cells against neoplastic cells, and cells infected with  
XX virus, bacterium, protozoan or fungus. CTL epitopes are identified based  
XX on the knowledge that such epitopes are, in fact, produced by the  
XX housekeeping proteasome system. Once identified, these epitopes, embodied  
XX as peptides, can be used to successfully immunise or induce therapeutic  
XX CTL responses against housekeeping proteasome expressing target cells in  
XX the host. The present sequence is that of a protein which is related to  
XX the method of the invention.  
XX Sequence 188 AA;  
Query Match 98.8%; Score 1000; DB 8; Length 188;  
Best Local Similarity 98.9%; Pred. No. 1.1e-99;  
Matches 186; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 MNGDDAFARRPTVGAQIPEKIQAFDDIAKYFSKEEWMKASEKIFVYMKRYEAMTK 60  
Db 1 MNGDDAFARRPTVGAQIPEKIQAFDDIAKYFSKEEWMKASEKIFVYMKRYEAMTK 60  
QY 61 LGFKATLPFFMCMKRAEDFGNDLNDPNRGNQVERPQMTFGRLOGISPKIMPKPAEAG 120  
Db 61 LGFKATLPFFMCMKRAEDFGNDLNDPNRGNQVERPQMTFGRLOGISPKIMPKPAEAG 120  
QY 121 NDSSEVPEASGPONDGKELCPGKPTTSKIHRSRGPKGGEHAWTHRLRERKOLVYEEI 180  
Db 121 NDSSEVPEASGPONDGKELCPGKPTTSKIHRSRGPKGGEHAWTHRLRERKOLVYEEI 180  
QY 181 SDEPDEDE 188  
Db 181 SDEPDEDE 188  
RESULT 11  
AAR90677  
ID AAR90677 standard; protein; 188 AA.

XX AAR90677;  
XX 16-APR-1996 (first entry)  
XX Human X-chromosome SSX2 protein.  
XX Human; SSX2; synovial sarcoma; X-chromosome; breakpoint-2;  
XX ornithine-delta-aminotransferase; OATL2 gene; Xp11.2; SSX1;  
XX translocation; chromosome-18; 18q11.2; X-chromosome; SYT; gene fusion;  
XX SYT-SSX2; fusion protein; primer; PCR; polymerase chain reaction; probe;  
XX antibody; monoclonal antibody; humanised antibody; hybridisation;  
XX antisense; antitumour; recombinant vaccine; vaccinia virus; vector;  
XX cancer; diagnosis; therapy.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
XX Region 110..111  
XX FT /note="Breakpoint in SYT translocation"  
XX Region 111..188  
XX FT /note="Sequence present in SYT-SSX2 fusion"  
XX PN WO9602641-A2.  
XX 01-FEB-1996.  
XX PF 19-JUL-1995; 95WO-GB001704.  
XX 19-JUL-1994; 94GB-00014580.  
XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.  
XX Cooper CS, Gusterson BA;  
XX WPI: 1996-105904/11.  
XX N-PSDB; AAT11780.  
XX Diagnosing synovial sarcoma by detection of e.g. SYT and SSX1 sequences,  
XX based on X18 trans-location - by amplification of trans-location  
XX sequences or detection of the encoded protein.  
XX Claim 6; Fig 7; 55pp; English.  
XX The sequence is encoded by the SSX2 gene (synovial sarcoma X-chromosome  
XX breakpoint-2). The gene is at the location of a breakpoint at Xp11.2  
XX within an ornithine-delta-aminotransferase OATL2 region, associated with  
XX the translocation t(X;18) (p11.2;q11.2) found in human synovial sarcomas,  
XX involving joining of the SYT gene (encoding AAR90675) on chromosome-18 at  
XX 18q11.2 to SSX2. The fusion gene is then translated into an SYT-SSX2  
XX (AAR90678) fusion protein. Detection of the gene or its product either in  
XX an abnormal location or as a fusion may be used in diagnosis of synovial  
XX sarcoma, using primers, probes, humanised antibodies, monoclonal  
XX antibodies, etc. Antisense oligonucleotides and antibodies may also be  
XX used therapeutically, and antitumour recombinant vaccines may be  
XX constructed e.g. in a vaccinia virus vector  
XX Sequence 188 AA;  
Query Match 97.6%; Score 988; DB 2; Length 188;  
Best Local Similarity 97.9%; Pred. No. 1.6e-88;  
Matches 184; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MNGDDAFARRPTVGAQIPEKIQAFDDIAKYFSKEEWMKASEKIFVYMKRYEAMTK 60  
Db 1 MNGDDAFARRPTVGAQIPEKIQAFDDIAKYFSKEEWMKASEKIFVYMKRYEAMTK 60  
QY 61 LGFKATLPFFMCMKRAEDFGNDLNDPNRGNQVERPQMTFGRLOGISPKIMPKPAEAG 120  
Db 61 LGFKATLPFFMCMKRAEDFGNDLNDPNRGNQVERPQMTFGRLOGISPKIMPKPAEAG 120  
QY 121 NDSSEVPEASGPONDGKELCPGKPTTSKIHRSRGPKGGEHAWTHRLRERKOLVYEEI 180

Db 121 NDSEVPASGQNDGKELCPPGKPTTSKIHRSKPGRGEHAWTHRLRERKQLVIYBEI 180

QY 181 SDPEEDDE 188  
|||||

Db 181 SDPEEDDE 188

RESULT 12

ID ADQ10464 standard; protein; 207 AA.

XX ADQ10464;

XX 23-SEP-2004 (first entry)

XX Epitope liberation associated polynucleotide seqid 29.

XX immunostimulant; cytostatic; vaccine; tumour-associated antigen SSX-2;

KW SSX-2 antigen; epitope cluster; MHC receptor peptide binding cleft;

KW immunogenic composition; immune response; cancer; vaccine vector;

KW epitope liberation; human leukocyte antigen; HLA A2-specific CTL;

KW cytotoxic T lymphocyte; human; epitope liberation; ds.

XX Synthetic.

OS US2004132088-A1.

PN 08-JUL-2004.

PD 10-FEB-2004; 2004US-00777053.

PF 07-NOV-2001; 2001US-0336968P.

PR 07-NOV-2002; 2002US-00292413.

XX (SIMA/) SIMARD J J L.

PA (DIAM/) DIAMOND D C.

PA (QIUZ/) QIU Z.

PA (LEIX/) LEI X.

XX Simard JJJ, Diamond DC, Qiu Z, Lei X;

PI WPI; 2004-517003/49.

DR Novel nucleic acid encoding tumor-associated antigen SSX-2, useful in

PT inducing an immune response and in treating cancer.

XX Example 5; SEQ ID NO 29; 260pp; English.

XX The invention describes an isolated nucleic acid (I) comprising a reading

CC frame comprising a first sequence, where the first sequence encodes one

CC or more segments of tumour-associated antigen SSX-2, which comprises a

CC sequence of 188 amino acids (SEQ ID NO: 40), where the first sequence

CC does not encode the complete SSX-2 antigen, and where each segment

CC comprises an epitope cluster, the cluster comprising or encoding at least

CC two amino acid sequences having a known or predicted affinity for a same

CC MHC receptor peptide binding cleft. Also described are: an isolated

CC polypeptide comprising the amino acid sequence encoded in the reading

CC of (I). (I) is a nucleic acid encoding a tumour-associated antigen SSX-2

CC comprising a fully defined sequence of 188 amino acids (SEQ ID NO: 40).

CC The nucleic acid, the encoded antigen, and composition are useful in

CC inducing an immune response and in treating cancer. Expression cassettes

CC are used in vaccine vectors. This is the amino acid sequence of a

CC polynucleotide associated with epitope liberation.

XX Sequence 207 AA;

Query Match 90.0%; Score 911; DB 8; Length 207;

Best Local Similarity 98.8%; Pred. No. 6.4e-81;

Matches 169; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 15 AQIPEKIQAFDDIAKYFSKEWEKASEKIFYVMKRYEAMTKLGFKATLPPFMCK 74  
|||||

Db 2 AQIPEKIQAFDDIAKYFSKEWEKASEKIFYVMKRYEAMTKLGFKATLPPFMCK 61

QY 75 RAEDFQGNLDNDNPNRGVQVERPQMTFGRLOGISPKIMPKPAEGRGNDSEEVPEASGPQN 134  
|||||

Db 62 RAEDFQGNLDNDNPNRGVQVERPQMTFGRLOGISPKIMPKPAEGRGNDSEEVPEASGPQN 121  
|||||

QY 135 DGKELCPPGKPTTSKIHRSKPGRGEHAWTHRLRERKQLVIYBEISDPEE 185  
|||||

Db 122 DGKELCPPGKPTTSKIHRSKPGRGEHAWTHRLRERKQLVIYBEISDPTQ 172  
|||||

RESULT 13

ADQ10463

ID ADQ10463 standard; protein; 208 AA.

XX ADQ10463;

XX 23-SEP-2004 (first entry)

XX Epitope liberation associated polynucleotide seqid 28.

XX immunostimulant; cytostatic; vaccine; tumour-associated antigen SSX-2;

KW SSX-2 antigen; epitope cluster; MHC receptor peptide binding cleft;

KW immunogenic composition; immune response; cancer; vaccine vector;

KW epitope liberation; human leukocyte antigen; HLA A2-specific CTL;

KW cytotoxic T lymphocyte; human; epitope liberation; ds.

XX Synthetic.

OS US2004132088-A1.

PN 08-JUL-2004.

PD 10-FEB-2004; 2004US-00777053.

PF 07-NOV-2001; 2001US-0336968P.

PR 07-NOV-2002; 2002US-00292413.

XX (SIMA/) SIMARD J J L.

PA (DIAM/) DIAMOND D C.

PA (QIUZ/) QIU Z.

PA (LEIX/) LEI X.

XX Simard JJJ, Diamond DC, Qiu Z, Lei X;

PI WPI; 2004-517003/49.

DR Novel nucleic acid encoding tumor-associated antigen SSX-2, useful in

PT inducing an immune response and in treating cancer.

XX Example 5; SEQ ID NO 28; 260pp; English.

XX The invention describes an isolated nucleic acid (I) comprising a reading

CC frame comprising a first sequence, where the first sequence encodes one

CC or more segments of tumour-associated antigen SSX-2, which comprises a

CC sequence of 188 amino acids (SEQ ID NO: 40), where the first sequence

CC does not encode the complete SSX-2 antigen, and where each segment

CC comprises an epitope cluster, the cluster comprising or encoding at least

CC two amino acid sequences having a known or predicted affinity for a same

CC MHC receptor peptide binding cleft. Also described are: an isolated

CC polypeptide comprising the amino acid sequence encoded in the reading

CC of (I). (I) is a nucleic acid encoding a tumour-associated antigen SSX-2

CC comprising a fully defined sequence of 188 amino acids (SEQ ID NO: 40).

CC The nucleic acid, the encoded antigen, and composition are useful in

CC inducing an immune response and in treating cancer. Expression cassettes

CC are used in vaccine vectors. This is the amino acid sequence of a

CC polynucleotide associated with epitope liberation.

XX Sequence 208 AA;

Query Match 90.0%; Score 910.5; DB 8; Length 208;

Best Local Similarity 95.0%; Pred. No. 7.2e-81;



Matches	172;	Conservative	0;	Mismatches	0;	Indels	9;	Gaps	1;
Qy	12	TVG-----	AAIQPEKIQAFDDIAKYFSKEWEKMKASEKIFVYVMKEKYEAMTKLG	62					
Db	28	TVGLPSIPVHPIAQIPEKIQAFDDIAKYFSKEWEKMKASEKIFVYVMKEKYEAMTKLG	87						
Qy	63	FKATLPPMCMKRAEDFGNDLNDPNRGNQVERPQMTFGRLOGISPKIMPKKPAEAGND	122						
Db	88	FKATLPPMCMKRAEDFGNDLNDPNRGNQVERPQMTFGRLOGISPKIMPKKPAEAGND	147						
Qy	123	SEEVPEASGPONDGKELCPGCKPTTSEKIHRSRSGPKRGEHAWTHRLRERKOLVYEEISD	182						
Db	148	SEEVPEASGPONDGKELCPGCKPTTSEKIHRSRSGPKRGEHAWTHRLRERKOLVYEEISD	207						
Qy	183	P 183							
Db	208	P 208							
RESULT 14									
ADQ10460	ID	ADQ10460	standard; protein; 169 AA.						
XX	AC	ADQ10460;							
XX	AC								
XX	DT								
XX	DT	23-SEP-2004	(first entry)						
XX	XX		Human tumour-associated antigen SSX-2 epitope seqid 25.						
XX	XX		immunostimulant; cytostatic; vaccine; tumour-associated antigen SSX-2;						
KW	KW		SSX-2 antigen; epitope cluster; MHC receptor peptide binding cleft;						
KW	KW		immunogenic composition; immune response; cancer; vaccine vector;						
KW	KW		epitope liberation; human leukocyte antigen; HLA A2-specific CTL;						
KW	KW		cytotoxic T lymphocyte; human; synovial sarcoma X breakpoint 2; epitope.						
XX	XX								
OS	OS		Homo sapiens.						
XX	XX								
PN	PN	US2004132088-A1.							
XX	XX								
PD	PD	08-JUL-2004.							
XX	XX								
PF	PF	10-FEB-2004; 2004US-00777053.							
XX	XX								
PR	PR	07-NOV-2001; 2001US-0336968P.							
XX	XX								
PR	PR	07-NOV-2002; 2002US-00292413.							
XX	XX								
PA	PA	(SIMA/) SIMARD J J L.							
PA	PA	(DIAM/) DIAMOND D C.							
PA	PA	(QIUZ/) QIU Z.							
XX	XX	(LEIX/) LEI X.							
PI	PI	Simard JJu, Diamond DC, Qiu Z, Lei X;							
XX	XX								
DR	DR	WPI; 2004-517003/49.							
XX	XX								
PT	PT	Novel nucleic acid encoding tumor-associated antigen SSX-2, useful in							
PT	PT	inducing an immune response and in treating cancer.							
XX	XX								
PS	PS	Example 5; SEQ ID NO 25; 260pp; English.							
XX	XX								
CC	CC	The invention describes an isolated nucleic acid (I) comprising a reading							
CC	CC	frame comprising a first sequence, where the first sequence encodes one							
CC	CC	or more segments of tumour-associated antigen SSX-2, which comprises a							
CC	CC	sequence of 188 amino acids (SEQ ID NO: 40), where the first sequence							
CC	CC	does not encode the complete SSX-2 antigen, and where each segment							
CC	CC	comprises an epitope cluster, the cluster comprising or encoding at least							
CC	CC	two amino acid sequences having a known or predicted affinity for a same							
CC	CC	MHC receptor peptide binding cleft. Also described are: an isolated							
CC	CC	polypeptide comprising the amino acid sequence encoded in the reading							
CC	CC	frame; and an immunogenic composition comprising (I) or the polypeptide							
CC	CC	of (I). (I) is a nucleic acid encoding a tumour-associated antigen SSX-2							
CC	CC	comprising a fully defined sequence of 188 amino acids (SEQ ID NO: 40).							
CC	CC	The nucleic acid, the encoded antigen, and composition are useful in							

CC	inducing an immune response and in treating cancer. Expression cassettes are used in vaccine vectors. This is the amino acid sequence of a human tumour-associated antigen SSX-2 epitope used to identify a polypeptide suitable for epitope liberation.
XX	
XX	
SQ	Sequence 169 AA;
	Query Match           89.9%; Score 910; DB 8; Length 169; Best Local Similarity 100.0%; Pred. No. 6.2e-81; Matches 169; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY	15 AQIPEKIQKAFDDDTAKYFSKEEWEKKASEKIFVYMKRKYEAETKLGFKATLPPFMCNK 74                                                                                                                                                             DB 1 AQIPEKIQKAFDDDTAKYFSKEEWEKKASEKIFVYMKRKYEAETKLGFKATLPPFMCNK 60                                                                               QY 75 RAEDFQNGDLDNDNRNGQVERPOMTFGRLOGISPKIMPKPAPEGNDSEEVPEASGPQN 134                                                                               DB 61 RAEDFQNGDLDNDNRNGQVERPOMTFGRLOGISPKIMPKPAPEGNDSEEVPEASGPQN 120                                                                               QY 135 DGKELCPGPGPTTSEKIHERSGPKRGSHAWTHRLRRKQLVIYEISDP 183                                                                               DB 121 DGKELCPGPGPTTSEKIHERSGPKRGSHAWTHRLRRKQLVIYEISDP 169 
RESULT 15	
ADQ10461	ID ADQ10461 standard; protein; 245 AA.
XX	
AC	ADQ10461;
XX	
DT	23-SEP-2004 (first entry)
XX	
DE	Epitope liberation associated polynucleotide seqid 26.
XX	
KW	immunostimulant; cytostatic; vaccine; tumour-associated antigen SSX-2; SSX-2 antigen; epitope cluster; MHC receptor peptide binding cleft; immunogenic composition; immune response; cancer; vaccine vector; epitope liberation; human leukocyte antigen; HLA A2-specific CTL; cytotoxic T lymphocyte; human; epitope liberation; ds. Synthetic.
OS	
XX	
FN	US2004132088-A1.
XX	
PD	08-JUL-2004.
XX	
PF	10-FEB-2004; 2004US-00777053.
XX	
PR	07-NOV-2001; 2001US-0336968P.
PR	07-NOV-2002; 2002US-00292413.
XX	(SIMA/) SIMARD J J L.
PA	(DIAM/) DIAMOND D C.
PA	(QIUZ/) QIU Z.
PA	(LEIX/) LEI X.
XX	
PI	Simard JJL, Diamond DC, Qiu Z, Lei X;
XX	
DR	WPI; 2004-517003/49.
XX	
PT	Novel nucleic acid encoding tumor-associated antigen SSX-2, useful in inducing an immune response and in treating cancer.
FT	
XX	
PS	Example 5; SEQ ID NO 26; 260pp; English.
XX	
CC	The invention describes an isolated nucleic acid (I) comprising a reading frame comprising a first sequence, where the first sequence encodes one or more segments of tumour-associated antigen SSX-2, which comprises a sequence of 188 amino acids (SEQ ID NO: 40), where the first sequence does not encode the complete SSX-2 antigen, and where each segment CC comprises an epitope cluster, the cluster comprising or encoding at least two amino acid sequences having a known or predicted affinity for a same MHC receptor peptide binding cleft. Also described are: an isolated

CC polypeptide comprising the amino acid sequence encoded in the reading  
CC frame; and an immunogenic composition comprising (I) or the polypeptide  
CC of (1). (I) is a nucleic acid encoding a tumour-associated antigen SSX-2  
CC comprising a fully defined sequence of 188 amino acids (SEQ ID NO: 40).  
CC The nucleic acid, the encoded antigen, and composition are useful in  
CC inducing an immune response and in treating cancer. Expression cassettes  
CC are used in vaccine vectors. This is the amino acid sequence of a  
CC polynucleotide associated with epitope liberation.  
XX  
SQ Sequence 245 AA;  
  
Query Match 89.9%; Score 910; DB 8; Length 245;  
Best Local Similarity 100.0%; Pred. No. 9.9e-81;  
Matches 169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 15 AQIPEKIQAFDDIAKYFSKEWEKKKASEKIFYVYMKRKYEAAMTKLGFKATLPPFMCNK 74  
Db 77 AQIPEKIQAFDDIAKYFSKEWEKKKASEKIFYVYMKRKYEAAMTKLGFKATLPPFMCNK 136  
  
Qy 75 RAEDFQGNLNDPNRGNQVERPQMTFGRLOGISPKIMPKKPAEEGNDSEEVPEASGFQN 134  
Db 137 RAEDFQGNLNDPNRGNQVERPQMTFGRLOGISPKIMPKKPAEEGNDSEEVPEASGFQN 196  
  
Qy 135 DGKELCPGKPTTSEKIHRSQPKRGEHAWTHRLRERKQLVIYEISDP 183  
Db 197 DGKELCPGKPTTSEKIHRSQPKRGEHAWTHRLRERKQLVIYEISDP 245

. Search completed: June 20, 2005, 12:18:18  
Job time : 166 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2005, 12:15:40 ; Search time 161 Seconds  
(without alignments)  
448.374 Million cell updates/sec

Title: US-10-777-053-40  
Perfect score: 1012  
Sequence: 1 MNGDDAFARRPTVGAQIPEK.....RRKQLVIYEISDPEDDE 188

Scoring table: BLOSOM62  
Gapop 10.0 , Gapext 0.5

Searched: 1714042 seqs, 383979560 residues

Total number of hits satisfying chosen parameters: 1714042

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database :
- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
  - 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/ptodata/2/pubpaa/PCTU5\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
  - 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
  - 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
  - 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
  - 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
  - 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
  - 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
  - 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
  - 17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*
  - 18: /cgn2\_6/ptodata/2/pubpaa/US10F\_NEW\_PUB.pep.\*
  - 19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*
  - 20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
  - 21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
  - 22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1.	1012	100.0	188	10	US-09-849-602-28
2.	1012	100.0	188	14	US-10-207-655-85
3.	1012	100.0	188	15	US-10-026-066-2
4.	1012	100.0	188	15	US-10-117-937-3
5.	1012	100.0	188	16	US-10-777-053-40
6.	1012	100.0	188	16	US-10-657-022-3
7.	1012	100.0	188	16	US-10-837-217-40
8.	1012	100.0	188	17	US-10-937-794-2
9.	911	90.0	207	16	US-10-777-053-29
10	911	90.0	207	16	US-10-837-217-29
11	910.5	90.0	208	16	US-10-777-053-28

12	910.5	90.0	208	16	US-10-837-217-28	Sequence 28, Appl
13	910	89.9	169	16	US-10-777-053-25	Sequence 25, Appl
14	910	89.9	169	16	US-10-837-217-25	Sequence 25, Appl
15	910	89.9	245	16	US-10-777-053-26	Sequence 26, Appl
16	910	89.9	245	16	US-10-777-053-27	Sequence 27, Appl
17	910	89.9	245	16	US-10-837-217-26	Sequence 27, Appl
18	910	89.9	245	16	US-10-837-217-27	Sequence 27, Appl
19	845.5	83.5	223	17	US-10-937-794-4	Sequence 4, Appl
20	845.5	83.5	223	17	US-10-871-708-3	Sequence 3, Appl
21	813	80.3	169	17	US-10-937-794-6	Sequence 6, Appl
22	798	78.9	188	15	US-10-117-937-598	Sequence 598, App
23	798	78.9	188	15	US-10-260-708-77	Sequence 77, Appl
24	798	78.9	188	15	US-10-657-022-94	Sequence 94, Appl
25	772	76.3	188	15	US-10-260-708-71	Sequence 71, Appl
26	363.5	35.9	492	16	US-10-408-765A-1141	Sequence 1141, Ap
27	201	19.9	38	15	US-10-117-937-10	Sequence 10, Appl
28	201	19.9	38	16	US-10-657-022-10	Sequence 10, Appl
29	135	13.3	752	15	US-10-094-749-2599	Sequence 2599, Ap
30	127	12.5	22	17	US-10-937-794-19	Sequence 19, Appl
31	124	12.3	22	17	US-10-937-794-18	Sequence 18, Appl
32	123	12.2	22	17	US-10-937-794-17	Sequence 17, Appl
33	122	12.1	22	17	US-10-937-794-12	Sequence 12, Appl
34	122	12.1	22	17	US-10-937-794-20	Sequence 20, Appl
35	121	12.0	22	17	US-10-937-794-13	Sequence 13, Appl
36	119	11.8	831	16	US-10-479-435-7	Sequence 7, Appl
37	118	11.7	22	17	US-10-937-794-10	Sequence 10, Appl
38	116	11.5	22	17	US-10-937-794-9	Sequence 9, Appl
39	116	11.5	22	17	US-10-937-794-11	Sequence 11, Appl
40	116	11.5	22	17	US-10-937-794-15	Sequence 15, Appl
41	115	11.4	22	17	US-10-937-794-16	Sequence 16, Appl
42	114	11.3	22	17	US-10-937-794-7	Sequence 7, Appl
43	113	11.2	21	17	US-10-937-794-56	Sequence 56, Appl
44	111	11.0	22	17	US-10-937-794-8	Sequence 8, Appl
45	109	10.8	20	17	US-10-937-794-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1  
US-09-849-602-28  
; Sequence 28, Application US/09849602  
; Publication No. US20030165834A1  
; GENERAL INFORMATION:  
; APPLICANT: Scanlan, Matthew J.  
; APPLICANT: Old, Lloyd J.  
; APPLICANT: Stockert, Elisabeth  
; APPLICANT: Chen, Yao-Tseung  
; TITLE OF INVENTION: Colon Cancer Antigen Panel  
; FILE REFERENCE: L0461/7105 (JRV)  
; CURRENT APPLICATION NUMBER: US/09/849,602  
; CURRENT FILING DATE: 2001-05-04  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 28  
; LENGTH: 188  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-849-602-28

Query Match 100.0%; Score 1012; DB 10; Length 188;  
Best Local Similarity 100.0%; Pred. No. 2.3e-86;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MNGDDAFARRPTVGAQIPEKIQAFDDIAKYSKBEWEKMKASEKIFYVYMKRYEAMTK	60
Db	1	MNGDDAFARRPTVGAQIPEKIQAFDDIAKYSKBEWEKMKASEKIFYVYMKRYEAMTK	60
Qy	61	LGFKATLPPFMCNKRAEDFQGNLDNDPNRGNQVERPQMTFGRLOGISPKIMPKPAEEG	120
Db	61	LGFKATLPPFMCNKRAEDFQGNLDNDPNRGNQVERPQMTFGRLOGISPKIMPKPAEEG	120
Qy	121	NDSEVPEASGQNDGKELCPFGKPTTSKIHRSQPKGEHAWTHRLRERQQLVYIEI	180

```
Db      121 NDSEVPEASGPDNGKELCPGKPTTSEKIHRSQPKRGEHAWTHRLRERKQLVYBEI 180
Qy      181 SDPEEDE 188
Db      181 SDPEEDE 188
```

## RESULT 2

```
US-10-207-655-85
; Sequence 85, Application US/10207655
; Publication No. US2003018592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 85
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-85
```

```
Query Match      100.0%; Score 1012; DB 14; Length 188;
Best Local Similarity 100.0%; Pred. No. 2.3e-86;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MNGDDAFARRPTVGAQIPEKIQAFDDIAKYFSKEWEKMKASEKIFVYMKRKYEAMTK 60
Db      1 MNGDDAFARRPTVGAQIPEKIQAFDDIAKYFSKEWEKMKASEKIFVYMKRKYEAMTK 60

Qy      61 LGFKATLPPFMCNKRAEDFQGNLDNDPNRGNQVERPQMTFGRLOGISPKIMPKKPABEG 120
Db      61 LGFKATLPPFMCNKRAEDFQGNLDNDPNRGNQVERPQMTFGRLOGISPKIMPKKPABEG 120

Qy      121 NDSEVPEASGPDNGKELCPGKPTTSEKIHRSQPKRGEHAWTHRLRERKQLVYBEI 180
Db      121 NDSEVPEASGPDNGKELCPGKPTTSEKIHRSQPKRGEHAWTHRLRERKQLVYBEI 180

Qy      181 SDPEEDE 188
Db      181 SDPEEDE 188
```

## RESULT 3

```
US-10-026-066-2
; Sequence 2, Application US/10026066
; Publication No. US20030215425A1
; GENERAL INFORMATION:
; APPLICANT: Simard, John J. L.
; APPLICANT: Diamond, David C.
; TITLE OF INVENTION: EPTOPE SYNCHRONIZATION IN ANTIGEN
; FILE REFERENCE: PRESENTING CELLS
; FILE REFERENCE: CTLIMM.21CP1C
; CURRENT APPLICATION NUMBER: US/10/026,066
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 09/561,074
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 09/560,465
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 09/561,572
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 09/561,571
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: PCT/US01/13806
; PRIOR FILING DATE: 2001-04-27
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
```

```
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-026-066-2
```

```
Query Match      100.0%; Score 1012; DB 15; Length 188;
Best Local Similarity 100.0%; Pred. No. 2.3e-86;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MNGDDAFARRPTVGAQIPEKIQAFDDIAKYFSKEWEKMKASEKIFVYMKRKYEAMTK 60
Db      1 MNGDDAFARRPTVGAQIPEKIQAFDDIAKYFSKEWEKMKASEKIFVYMKRKYEAMTK 60

Qy      61 LGFKATLPPFMCNKRAEDFQGNLDNDPNRGNQVERPQMTFGRLOGISPKIMPKKPABEG 120
Db      61 LGFKATLPPFMCNKRAEDFQGNLDNDPNRGNQVERPQMTFGRLOGISPKIMPKKPABEG 120

Qy      121 NDSEVPEASGPDNGKELCPGKPTTSEKIHRSQPKRGEHAWTHRLRERKQLVYBEI 180
Db      121 NDSEVPEASGPDNGKELCPGKPTTSEKIHRSQPKRGEHAWTHRLRERKQLVYBEI 180

Qy      181 SDPEEDE 188
Db      181 SDPEEDE 188
```

## RESULT 4

```
US-10-117-937-3
; Sequence 3, Application US/10117937
; Publication No. US20030220239A1
; GENERAL INFORMATION:
; APPLICANT: CTL IMMUNO THERAPIES CORP.
; APPLICANT: SIMARD, John, J.L.
; APPLICANT: DIAMOND, David, C.
; APPLICANT: LIU, Liping
; APPLICANT: XIE, Zhigong
; FILE OF INVENTION: EPTOPE SEQUENCES
; FILE REFERENCE: CTLIMM.027A
; CURRENT APPLICATION NUMBER: US/10/117,937
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: US 60/282,211
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: US 60/337,017
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/363,210
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 602
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-937-3
```

```
Query Match      100.0%; Score 1012; DB 15; Length 188;
Best Local Similarity 100.0%; Pred. No. 2.3e-86;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MNGDDAFARRPTVGAQIPEKIQAFDDIAKYFSKEWEKMKASEKIFVYMKRKYEAMTK 60
Db      1 MNGDDAFARRPTVGAQIPEKIQAFDDIAKYFSKEWEKMKASEKIFVYMKRKYEAMTK 60

Qy      61 LGFKATLPPFMCNKRAEDFQGNLDNDPNRGNQVERPQMTFGRLOGISPKIMPKKPABEG 120
Db      61 LGFKATLPPFMCNKRAEDFQGNLDNDPNRGNQVERPQMTFGRLOGISPKIMPKKPABEG 120

Qy      121 NDSEVPEASGPDNGKELCPGKPTTSEKIHRSQPKRGEHAWTHRLRERKQLVYBEI 180
Db      121 NDSEVPEASGPDNGKELCPGKPTTSEKIHRSQPKRGEHAWTHRLRERKQLVYBEI 180

Qy      181 SDPEEDE 188
Db      181 SDPEEDE 188
```

RESULT 5  
US-10-777-053-40  
; Sequence 40, Application US/10777053  
; Publication No. US20040132088A1  
; GENERAL INFORMATION:  
; APPLICANT: Simard, John J. L.  
; APPLICANT: Diamond, David C.  
; APPLICANT: Qiu, Zhiyong  
; APPLICANT: Lei, Xiang-Dong  
; TITLE OF INVENTION: EXPRESSION VECTORS ENCODING EPITOPES OF  
; TARGET-ASSOCIATED ANTIGENS AND METHODS FOR THEIR DESIGN  
; FILE REFERENCE: MANK.022C1  
; CURRENT APPLICATION NUMBER: US/10/777,053  
; CURRENT FILING DATE: 2004-02-10  
; PRIOR APPLICATION NUMBER: 10/292,413  
; PRIOR FILING DATE: 2002-11-07  
; PRIOR APPLICATION NUMBER: 60/336,968  
; PRIOR FILING DATE: 2001-11-07  
; NUMBER OF SEQ ID NOS: 979  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 40  
; LENGTH: 188  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-777-053-40

Query Match 100.0%; Score 1012; DB 16; Length 188;  
Best Local Similarity 100.0%; Pred. No. 2.3e-86;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MNGDDAFARRPTVGAQIPEKIQAFDDIAKYFSKEWEKMKASEKIFYYVMKRYEAMTK 60  
DB 1 MNGDDAFARRPTVGAQIPEKIQAFDDIAKYFSKEWEKMKASEKIFYYVMKRYEAMTK 60  
QY 61 LGFKATLPPFMCNKAEDFQGNLDNDNRGNQVERPQMTFGRLOGISPKIMPKPAEAG 120  
DB 61 LGFKATLPPFMCNKAEDFQGNLDNDNRGNQVERPQMTFGRLOGISPKIMPKPAEAG 120  
QY 121 NDSEVPASGPNQDGKELCPGKPTTSEKIHSGPKRGEHAWTHRLRERKQVLYEEI 180  
DB 121 NDSEVPASGPNQDGKELCPGKPTTSEKIHSGPKRGEHAWTHRLRERKQVLYEEI 180  
QY 181 SDPEEDE 188  
DB 181 SDPEEDE 188

RESULT 6  
US-10-657-022-3  
; Sequence 3, Application US/10657022  
; Publication No. US20040180354A1  
; GENERAL INFORMATION:  
; APPLICANT: Simard, John J. L.  
; APPLICANT: Diamond, David C.  
; APPLICANT: Liu, Liping  
; APPLICANT: Liu, Zheng  
; TITLE OF INVENTION: EPITOPE SEQUENCES  
; CURRENT APPLICATION NUMBER: US/10/657,022  
; CURRENT FILING DATE: 2003-09-04  
; PRIOR APPLICATION NUMBER: 60/409123  
; PRIOR FILING DATE: 2002-09-06  
; NUMBER OF SEQ ID NOS: 610  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 188  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-657-022-3

Query Match 100.0%; Score 1012; DB 16; Length 188;

Best Local Similarity 100.0%; Pred. No. 2.3e-86;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MNGDDAFARRPTVGAQIPEKIQAFDDIAKYFSKEWEKMKASEKIFYYVMKRYEAMTK 60  
DB 1 MNGDDAFARRPTVGAQIPEKIQAFDDIAKYFSKEWEKMKASEKIFYYVMKRYEAMTK 60  
QY 61 LGFKATLPPFMCNKAEDFQGNLDNDNRGNQVERPQMTFGRLOGISPKIMPKPAEAG 120  
DB 61 LGFKATLPPFMCNKAEDFQGNLDNDNRGNQVERPQMTFGRLOGISPKIMPKPAEAG 120  
QY 121 NDSEVPASGPNQDGKELCPGKPTTSEKIHSGPKRGEHAWTHRLRERKQVLYEEI 180  
DB 121 NDSEVPASGPNQDGKELCPGKPTTSEKIHSGPKRGEHAWTHRLRERKQVLYEEI 180  
QY 181 SDPEEDE 188  
DB 181 SDPEEDE 188

RESULT 7  
US-10-837-217-40  
; Sequence 40, Application US/10837217  
; Publication No. US20040203051A1  
; GENERAL INFORMATION:  
; APPLICANT: Simard, John J. L.  
; APPLICANT: Diamond, David C.  
; APPLICANT: Qiu, Zhiyong  
; APPLICANT: Lei, Xiang-Dong  
; TITLE OF INVENTION: EXPRESSION VECTORS ENCODING EPITOPES OF  
; TARGET-ASSOCIATED ANTIGENS AND METHODS FOR THEIR DESIGN  
; FILE REFERENCE: MANK.022C2  
; CURRENT APPLICATION NUMBER: US/10/837,217  
; CURRENT FILING DATE: 2004-04-30  
; PRIOR APPLICATION NUMBER: 10/292,413  
; PRIOR FILING DATE: 2002-11-07  
; PRIOR APPLICATION NUMBER: 60/336,968  
; PRIOR FILING DATE: 2001-11-07  
; NUMBER OF SEQ ID NOS: 979  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 40  
; LENGTH: 188  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-837-217-40

Query Match 100.0%; Score 1012; DB 16; Length 188;  
Best Local Similarity 100.0%; Pred. No. 2.3e-86;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MNGDDAFARRPTVGAQIPEKIQAFDDIAKYFSKEWEKMKASEKIFYYVMKRYEAMTK 60  
DB 1 MNGDDAFARRPTVGAQIPEKIQAFDDIAKYFSKEWEKMKASEKIFYYVMKRYEAMTK 60  
QY 61 LGFKATLPPFMCNKAEDFQGNLDNDNRGNQVERPQMTFGRLOGISPKIMPKPAEAG 120  
DB 61 LGFKATLPPFMCNKAEDFQGNLDNDNRGNQVERPQMTFGRLOGISPKIMPKPAEAG 120  
QY 121 NDSEVPASGPNQDGKELCPGKPTTSEKIHSGPKRGEHAWTHRLRERKQVLYEEI 180  
DB 121 NDSEVPASGPNQDGKELCPGKPTTSEKIHSGPKRGEHAWTHRLRERKQVLYEEI 180  
QY 181 SDPEEDE 188  
DB 181 SDPEEDE 188

RESULT 8  
US-10-937-794-2  
; Sequence 2, Application US/10937794  
; Publication No. US20050079553A1  
; GENERAL INFORMATION:  
; APPLICANT: Ayyoub, Maha

APPLICANT: Valmori, Danila  
; TITLE OF INVENTION: SSX-2 PEPTIDES PRESENTED BY HLA CLASS II MOLECULES  
; FILE REFERENCE: L0461.70162U00  
; CURRENT APPLICATION NUMBER: US/10/937,794  
; CURRENT FILING DATE: 2004-09-09  
; PRIOR APPLICATION NUMBER: US 10/779,568  
; PRIOR FILING DATE: 2004-02-13  
; PRIOR APPLICATION NUMBER: US 09/408,036  
; PRIOR FILING DATE: 1999-09-29  
; NUMBER OF SEQ ID NOS: 86  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 2  
; LENGTH: 188  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-937-794-2

Query Match 100.0%; Score 1012; DB 17; Length 188;  
Best Local Similarity 100.0%; Pred. No. 2.3e-86;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNGDDAFARRPTVGAQIPEKIQAFDDIAKYFSKEWEKMKASEKIFVYVMKRYEAMTK 60  
DB 1 MNGDDAFARRPTVGAQIPEKIQAFDDIAKYFSKEWEKMKASEKIFVYVMKRYEAMTK 60

QY 61 LGFKATLPPFMCKRAEDFQGNLDLNDPNRGNQVERPQMTFGRLOGISPKIMPKPAESEG 120  
DB 61 LGFKATLPPFMCKRAEDFQGNLDLNDPNRGNQVERPQMTFGRLOGISPKIMPKPAESEG 120

QY 121 NDSEVPASGPNQDKELCPGKPTTSEKIHRSRGRGEHAWTHRLRERKQLVYEI 180  
DB 121 NDSEVPASGPNQDKELCPGKPTTSEKIHRSRGRGEHAWTHRLRERKQLVYEI 180

QY 181 SDPEEDDE 188  
DB 181 SDPEEDDE 188

RESULT 9  
US-10-777-053-29  
; Sequence 29, Application US/10777053  
; Publication No. US20040132088A1  
; GENERAL INFORMATION:  
; APPLICANT: Simard, John J. L.  
; APPLICANT: Diamond, David C.  
; APPLICANT: Qiu, Zhiyong  
; APPLICANT: Lei, Xiang-Dong  
; TITLE OF INVENTION: EXPRESSION VECTORS ENCODING EPITOPES OF  
; FILE REFERENCE: MANNK.022C1  
; CURRENT APPLICATION NUMBER: US/10/777,053  
; CURRENT FILING DATE: 2004-02-10  
; PRIOR APPLICATION NUMBER: 10/292,413  
; PRIOR FILING DATE: 2002-11-07  
; PRIOR APPLICATION NUMBER: 60/336,968  
; PRIOR FILING DATE: 2001-11-07  
; NUMBER OF SEQ ID NOS: 979  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 29  
; LENGTH: 207  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: CTL54 expression product  
US-10-777-053-29

Query Match 90.0%; Score 911; DB 16; Length 207;  
Best Local Similarity 98.8%; Pred. No. 6.9e-77;  
Matches 169; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 15 AQIPEKIQAFDDIAKYFSKEWEKMKASEKIFVYVMKRYEAMTKLGFKATLPPFMCKN 74  
DB 2 AQIPEKIQAFDDIAKYFSKEWEKMKASEKIFVYVMKRYEAMTKLGFKATLPPFMCKN 61

QY 75 RAEDFQGNLDLNDPNRGNQVERPQMTFGRLOGISPKIMPKPAESEGNDSEVPASGPON 134  
DB 62 RAEDFQGNLDLNDPNRGNQVERPQMTFGRLOGISPKIMPKPAESEGNDSEVPASGPON 121

QY 135 DGKELCPPGKPTTSEKIHRSRGRGEHAWTHRLRERKQLVYEISDPEE 185  
DB 122 DGKELCPPGKPTTSEKIHRSRGRGEHAWTHRLRERKQLVYEISDPTQ 172

RESULT 11  
US-10-777-053-28  
; Sequence 28, Application US/10777053  
; Publication No. US20040132088A1  
; GENERAL INFORMATION:  
; APPLICANT: Simard, John J. L.  
; APPLICANT: Diamond, David C.  
; APPLICANT: Qiu, Zhiyong  
; APPLICANT: Lei, Xiang-Dong  
; TITLE OF INVENTION: EXPRESSION VECTORS ENCODING EPITOPES OF  
; FILE REFERENCE: MANNK.022C1  
; CURRENT APPLICATION NUMBER: US/10/777,053  
; CURRENT FILING DATE: 2004-02-10  
; PRIOR APPLICATION NUMBER: 10/292,413  
; PRIOR FILING DATE: 2002-11-07  
; PRIOR APPLICATION NUMBER: 60/336,968  
; PRIOR FILING DATE: 2001-11-07

QY 75 RAEDFQGNLDLNDPNRGNQVERPQMTFGRLOGISPKIMPKPAESEGNDSEVPASGPON 134  
DB 62 RAEDFQGNLDLNDPNRGNQVERPQMTFGRLOGISPKIMPKPAESEGNDSEVPASGPON 121

QY 135 DGKELCPPGKPTTSEKIHRSRGRGEHAWTHRLRERKQLVYEISDPEE 185  
DB 122 DGKELCPPGKPTTSEKIHRSRGRGEHAWTHRLRERKQLVYEISDPTQ 172

RESULT 10  
US-10-837-217-29  
; Sequence 29, Application US/10837217  
; Publication No. US20040203051A1  
; GENERAL INFORMATION:  
; APPLICANT: Simard, John J. L.  
; APPLICANT: Diamond, David C.  
; APPLICANT: Qiu, Zhiyong  
; APPLICANT: Lei, Xiang-Dong  
; TITLE OF INVENTION: EXPRESSION VECTORS ENCODING EPITOPES OF  
; FILE REFERENCE: MANNK.022C2  
; CURRENT APPLICATION NUMBER: US/10/837,217  
; CURRENT FILING DATE: 2004-04-30  
; PRIOR APPLICATION NUMBER: 10/292,413  
; PRIOR FILING DATE: 2002-11-07  
; PRIOR APPLICATION NUMBER: 60/336,968  
; PRIOR FILING DATE: 2001-11-07  
; NUMBER OF SEQ ID NOS: 979  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 29  
; LENGTH: 207  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: CTL54 expression product  
US-10-837-217-29

Query Match 90.0%; Score 911; DB 16; Length 207;  
Best Local Similarity 98.8%; Pred. No. 6.9e-77;  
Matches 169; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 15 AQIPEKIQAFDDIAKYFSKEWEKMKASEKIFVYVMKRYEAMTKLGFKATLPPFMCKN 74  
DB 2 AQIPEKIQAFDDIAKYFSKEWEKMKASEKIFVYVMKRYEAMTKLGFKATLPPFMCKN 61

QY 75 RAEDFQGNLDLNDPNRGNQVERPQMTFGRLOGISPKIMPKPAESEGNDSEVPASGPON 134  
DB 62 RAEDFQGNLDLNDPNRGNQVERPQMTFGRLOGISPKIMPKPAESEGNDSEVPASGPON 121

QY 135 DGKELCPPGKPTTSEKIHRSRGRGEHAWTHRLRERKQLVYEISDPEE 185  
DB 122 DGKELCPPGKPTTSEKIHRSRGRGEHAWTHRLRERKQLVYEISDPTQ 172

RESULT 11  
US-10-777-053-28  
; Sequence 28, Application US/10777053  
; Publication No. US20040132088A1  
; GENERAL INFORMATION:  
; APPLICANT: Simard, John J. L.  
; APPLICANT: Diamond, David C.  
; APPLICANT: Qiu, Zhiyong  
; APPLICANT: Lei, Xiang-Dong  
; TITLE OF INVENTION: EXPRESSION VECTORS ENCODING EPITOPES OF  
; FILE REFERENCE: MANNK.022C1  
; CURRENT APPLICATION NUMBER: US/10/777,053  
; CURRENT FILING DATE: 2004-02-10  
; PRIOR APPLICATION NUMBER: 10/292,413  
; PRIOR FILING DATE: 2002-11-07  
; PRIOR APPLICATION NUMBER: 60/336,968  
; PRIOR FILING DATE: 2001-11-07

```

; NUMBER OF SEQ ID NOS: 979
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: CTL53 expression product
US-10-777-053-28

Query Match
Best Local Similarity 90.0%; Score 910.5; DB 16; Length 208;
Matches 172; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

QY 12 TVG-----AQIPEKIQAADDIAKFSKEWEKMAKSEKIFVYMKRYEAMTKLG 62
DB 28 TVGLSPVHPHQAIPKIQKAFDDIAKFSKEWEKMAKSEKIFVYMKRYEAMTKLG 87
QY 63 FKATLPPFNCNRAEDFQGNLDNDNRGNQVERPQMTFGRLOGISPKIMPKPAEAGND 122
DB 88 FKATLPPFNCNRAEDFQGNLDNDNRGNQVERPQMTFGRLOGISPKIMPKPAEAGND 147
QY 123 SEEVPEASGPQDQKELCPGKPTTSEKIHRSRGPGRGEHAWTHRLRERKQLVYEEISD 182
DB 148 SEEVPEASGPQDQKELCPGKPTTSEKIHRSRGPGRGEHAWTHRLRERKQLVYEEISD 207
QY 183 P 183
DB 208 P 208

RESULT 12
US-10-837-217-28
; Sequence 28, Application US/10837217
; Publication No. US20040203051A1
; GENERAL INFORMATION:
; APPLICANT: Simard, John J. L.
; APPLICANT: Diamond, David C.
; APPLICANT: Qiu, Zhiyong
; APPLICANT: Lei, Xiang-Dong
; TITLE OF INVENTION: EXPRESSION VECTORS ENCODING EPITOPES OF
; FILE REFERENCE: TARGET-ASSOCIATED ANTIGENS AND METHODS FOR THEIR DESIGN
; CURRENT APPLICATION NUMBER: US/10/837,217
; PRIOR FILING DATE: 2004-04-30
; PRIOR APPLICATION NUMBER: 10/292,413
; PRIOR FILING DATE: 2002-11-07
; PRIOR APPLICATION NUMBER: 60/336,968
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 979
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: CTL53 expression product
US-10-837-217-28

Query Match
Best Local Similarity 90.0%; Score 910.5; DB 16; Length 208;
Matches 172; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

QY 12 TVG-----AQIPEKIQAADDIAKFSKEWEKMAKSEKIFVYMKRYEAMTKLG 62
DB 28 TVGLSPVHPHQAIPKIQKAFDDIAKFSKEWEKMAKSEKIFVYMKRYEAMTKLG 87
QY 63 FKATLPPFNCNRAEDFQGNLDNDNRGNQVERPQMTFGRLOGISPKIMPKPAEAGND 122
DB 88 FKATLPPFNCNRAEDFQGNLDNDNRGNQVERPQMTFGRLOGISPKIMPKPAEAGND 147
QY 123 SEEVPEASGPQDQKELCPGKPTTSEKIHRSRGPGRGEHAWTHRLRERKQLVYEEISD 182
DB 148 SEEVPEASGPQDQKELCPGKPTTSEKIHRSRGPGRGEHAWTHRLRERKQLVYEEISD 207
QY 183 P 183
DB 208 P 208

RESULT 13
US-10-777-053-25
; Sequence 25, Application US/10777053
; Publication No. US20040132088A1
; GENERAL INFORMATION:
; APPLICANT: Simard, John J. L.
; APPLICANT: Diamond, David C.
; APPLICANT: Qiu, Zhiyong
; APPLICANT: Lei, Xiang-Dong
; TITLE OF INVENTION: EXPRESSION VECTORS ENCODING EPITOPES OF
; FILE REFERENCE: TARGET-ASSOCIATED ANTIGENS AND METHODS FOR THEIR DESIGN
; CURRENT APPLICATION NUMBER: US/10/777,053
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: 10/292,413
; PRIOR FILING DATE: 2002-11-07
; PRIOR APPLICATION NUMBER: 60/336,968
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 979
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-777-053-25

Query Match
Best Local Similarity 89.9%; Score 910; DB 16; Length 169;
Matches 169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 AQIPEKIQAADDIAKFSKEWEKMAKSEKIFVYMKRYEAMTKLGKATLPPFMCNK 74
DB 1 AQIPEKIQAADDIAKFSKEWEKMAKSEKIFVYMKRYEAMTKLGKATLPPFMCNK 60
QY 75 RAEDFQGNLDNDNRGNQVERPQMTFGRLOGISPKIMPKPAEAGNDSEEVPEASGPON 134
DB 61 RAEDFQGNLDNDNRGNQVERPQMTFGRLOGISPKIMPKPAEAGNDSEEVPEASGPON 120
QY 135 DGKELCPGKPTTSEKIHRSRGPGRGEHAWTHRLRERKQLVYEEISDP 183
DB 121 DGKELCPGKPTTSEKIHRSRGPGRGEHAWTHRLRERKQLVYEEISDP 169

RESULT 14
US-10-837-217-25
; Sequence 25, Application US/10837217
; Publication No. US20040203051A1
; GENERAL INFORMATION:
; APPLICANT: Simard, John J. L.
; APPLICANT: Diamond, David C.
; APPLICANT: Qiu, Zhiyong
; APPLICANT: Lei, Xiang-Dong
; TITLE OF INVENTION: EXPRESSION VECTORS ENCODING EPITOPES OF
; FILE REFERENCE: TARGET-ASSOCIATED ANTIGENS AND METHODS FOR THEIR DESIGN
; CURRENT APPLICATION NUMBER: US/10/837,217
; CURRENT FILING DATE: 2004-04-30
; PRIOR APPLICATION NUMBER: 10/292,413
; PRIOR FILING DATE: 2002-11-07
; PRIOR APPLICATION NUMBER: 60/336,968
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 979
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 169
; TYPE: PRT
```

```
; ORGANISM: Homo Sapien
US-10-837-217-25

Query Match      89.9%; Score 910; DB 16; Length 169;
Best Local Similarity 100.0%; Pred. No. 6.6e-77;
Matches 169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 AQIPEKIQAFDDIAKYFSKEWEKWKASEKIFYVYMKRKYEAMTKLGFKATLPPFMCNK 74
   |||||
Db 1 AQIPEKIQAFDDIAKYFSKEWEKWKASEKIFYVYMKRKYEAMTKLGFKATLPPFMCNK 60
   |||||

QY 75 RAEDFGQNDLNDNRGNQVERPQMTFGRLOGISPKIMPKPABEGNDSEEVPEASGPON 134
   |||||
Db 61 RAEDFGQNDLNDNRGNQVERPQMTFGRLOGISPKIMPKPABEGNDSEEVPEASGPON 120
   |||||

QY 135 DGKELCPPGKPTTSEKIHRSQPKRGEHAWTHRLRERKQLVIYEISDP 183
   |||||
Db 121 DGKELCPPGKPTTSEKIHRSQPKRGEHAWTHRLRERKQLVIYEISDP 169
   |||||

RESULT 15
US-10-777-053-26
; Sequence 26, Application US/10777053
; Publication No. US20040132068A1
; GENERAL INFORMATION:
; APPLICANT: Simard, John J. L.
; APPLICANT: Diamond, David C.
; APPLICANT: Qiu, Zhiyong
; APPLICANT: Lei, Xiang-Dong
; TITLE OF INVENTION: EXPRESSION VECTORS ENCODING EPITOPES OF
; FILE REFERENCE: TARGET-ASSOCIATED ANTIGENS AND METHODS FOR THEIR DESIGN
; CURRENT APPLICATION NUMBER: US/10/777,053
; PRIOR FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: 10/292,413
; PRIOR FILING DATE: 2002-11-07
; PRIOR APPLICATION NUMBER: 60/336,968
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 979
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: CTLs1/PCBP expression product
US-10-777-053-26

Query Match      89.9%; Score 910; DB 16; Length 245;
Best Local Similarity 100.0%; Pred. No. 1.1e-76;
Matches 169; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 AQIPEKIQAFDDIAKYFSKEWEKWKASEKIFYVYMKRKYEAMTKLGFKATLPPFMCNK 74
   |||||
Db 77 AQIPEKIQAFDDIAKYFSKEWEKWKASEKIFYVYMKRKYEAMTKLGFKATLPPFMCNK 136
   |||||

QY 75 RAEDFGQNDLNDNRGNQVERPQMTFGRLOGISPKIMPKPABEGNDSEEVPEASGPON 134
   |||||
Db 137 RAEDFGQNDLNDNRGNQVERPQMTFGRLOGISPKIMPKPABEGNDSEEVPEASGPON 196
   |||||

QY 135 DGKELCPPGKPTTSEKIHRSQPKRGEHAWTHRLRERKQLVIYEISDP 183
   |||||
Db 197 DGKELCPPGKPTTSEKIHRSQPKRGEHAWTHRLRERKQLVIYEISDP 245
   |||||
```

Search completed: June 20, 2005, 12:25:40  
Job time : 163 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 20, 2005, 12:07:19 ; Search time 42 Seconds  
(without alignments)  
334.143 Million cell updates/sec

Title: US-10-777-053-40

Perfect score: 1012

Sequence: 1 MNGDDAFARRPTVGAQIPEK.....RERKQLVIYEEISDPEEDDE 188

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/prodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/prodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/prodata/1/iaa/PCUTUS\_COMB.pep.\*
- 6: /cgn2\_6/prodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1012	100.0	188	4	US-09-392-714-27
2	1012	100.0	188	4	US-09-849-602-28
3	798	78.9	188	4	US-09-392-714-28
4	119	11.8	675	1	US-08-317-522A-9
5	119	11.8	675	1	US-08-439-818A-9
6	119	11.8	675	2	US-08-751-965-9
7	119	11.8	675	2	US-08-738-975-9
8	119	11.8	675	2	US-08-728-626-9
9	119	11.8	675	3	US-08-808-599A-9
10	113	11.2	221	4	US-09-248-796A-16638
11	93.5	9.2	810	4	US-09-538-092-1275
12	87.5	8.6	208	4	US-09-538-092-1018
13	87.5	8.6	208	4	US-09-214-881A-2
14	87.5	8.6	320	4	US-09-949-016-10728
15	86.5	8.5	654	4	US-09-538-092-1193
16	86.5	8.5	854	4	US-09-949-016-6357
17	86	8.5	828	4	US-09-248-796A-20746
18	85.5	8.4	96	4	US-09-621-976-4810
19	85.5	8.4	118	4	US-09-489-039A-9797
20	85.5	8.4	506	2	US-08-820-170A-19
21	85.5	8.4	506	3	US-09-055-699-19
22	85.5	8.4	506	3	US-09-273-565-19
23	85.5	8.4	506	3	US-09-565-538-19
24	85.5	8.4	506	3	US-09-661-468-19
25	85.5	8.4	506	4	US-09-976-165-19
26	85.5	8.4	506	4	US-09-538-092-1381
27	85.5	8.4	663	4	US-09-949-016-7484

28 85.5 8.4 2404 4 US-09-949-016-11690 Sequence 11690, A  
29 85.5 8.4 2404 4 US-09-949-016-11691 Sequence 11691, A  
30 85.5 8.4 2547 3 US-09-538-489-35 Sequence 35, Appl  
31 85.5 8.4 2547 4 US-09-538-092-1374 Sequence 1374, Ap  
32 85 8.4 209 4 US-09-214-881A-6 Sequence 6, Appli  
33 85 8.4 671 3 US-09-121-321-16 Sequence 16, Appl  
34 85 8.4 671 3 US-08-933-803A-16 Sequence 16, Appl  
35 85 8.4 1167 2 US-08-589-756-2 Sequence 2, Appli  
36 85 8.4 1167 3 US-09-206-800-2 Sequence 2, Appli  
37 85 8.4 1167 3 US-09-206-898-2 Sequence 2, Appli  
38 84 8.3 15 4 US-09-408-036B-33 Sequence 33, Appl  
39 84 8.3 889 4 US-09-902-540-15702 Sequence 15702, A  
40 84 8.3 1647 4 US-09-824-574-4 Sequence 4, Appli  
41 84 8.3 1647 4 US-09-538-092-1172 Sequence 1172, Ap  
42 83.5 8.3 567 4 US-09-205-258-573 Sequence 573, App  
43 83.5 8.3 1341 4 US-09-949-016-6890 Sequence 6890, Ap  
44 83.5 8.3 1344 4 US-09-949-016-10925 Sequence 10925, A  
45 83.5 8.3 1829 3 US-09-157-420-1 Sequence 1, Appli

#### ALIGNMENTS

##### RESULT 1

US-09-392-714-27

; Sequence 27, Application US/09392714A

; Patent No. 6686147

; GENERAL INFORMATION:

; APPLICANT: Scanlan, Matthew J.

; APPLICANT: Gure, Ali O.

; APPLICANT: Williamson, Barbara

; APPLICANT: Chen, Yao-Tseng

; APPLICANT: Old, Lloyd J.

; TITLE OF INVENTION: Cancer Associated Antigens and Uses

; TITLE OF INVENTION: Therefor

; FILE REFERENCE: L0461/7062

; CURRENT APPLICATION NUMBER: US/09/392,714A

; EARLIER APPLICATION NUMBER: PCT/US98/14679

; EARLIER FILING DATE: 1998-07-15

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 27

; LENGTH: 188

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-392-714-27

Query Match 100.0%; Score 1012; DB 4; Length 188;  
Best Local Similarity 100.0%; Fred. No. 6.7e-104;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNGDDAFARRPTVGAQIPEKI QKAFDDIAKYFSKEWEKMKASEKIFVYVMKRYEAMTK 60

Db 1 MNGDDAFARRPTVGAQIPEKI QKAFDDIAKYFSKEWEKMKASEKIFVYVMKRYEAMTK 60

QY 61 LGFKATLPPFCMCKRAEDFGNDLNDPNRGNOVERPQMTFGRLOGISPKIMPKPAEEG 120

Db 61 LGFKATLPPFCMCKRAEDFGNDLNDPNRGNOVERPQMTFGRLOGISPKIMPKPAEEG 120

QY 121 NDSEVPEASGQNDGKELCPGKPTTSEKIHRSCKGEHAWTHRLRERKQLVIYEEI 180

Db 121 NDSEVPEASGQNDGKELCPGKPTTSEKIHRSCKGEHAWTHRLRERKQLVIYEEI 180

QY 181 SDPEEDDE 188

Db 181 SDPEEDDE 188

##### RESULT 2

US-09-849-602-28

; Sequence 28, Application US/09849602

; Patent No. 6794501

Db	61	LGFKVTLPPFMRKRAADFGNDNRHNRQVERPQMTFGSLQRIFPKIMPKKPAAEE	120
Qy	121	NDSEEVPEASGPONDGKELCPPGKPTTSEKIHERSGPKRGHEAWTHRLRERKQLVIYEEI	180
Db	121	NGLKEVPESAGPONDGKLCFPGNPSTLEKINKTSQPKRGKHAWTHRLRERKQLVVYEEI	180
Qy	181	SDPEEDEDE 188	
Db	181	SDPEEDEDE 188	
RESULT 4			
US-08-317-522A-9			
; Sequence 9, Application US/08317522A			
; Patent No. 559918			
; GENERAL INFORMATION:			
; APPLICANT: Fukuda, Michiko N.			
; TITLE OF INVENTION: Trophinin and Trophinin-Assisting			
; TITLE OF INVENTION: Proteins			
; NUMBER OF SEQUENCES: 13			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Campbell and Flores			
; STREET: 4370 La Jolla Village Drive, Suite 700			
; CITY: San Diego			
; STATE: California			
; COUNTRY: USA			
; ZIP: 92122			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.25			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/317,522A			
; FILING DATE: 04-OCT-1994			
; CLASSIFICATION: 536			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Campbell, Cathryn A.			
; REGISTRATION NUMBER: 31,815			
; REFERENCE/DOCKET NUMBER: P-LA 9991			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (619) 535-9001			
; TELEFAX: (619) 535-8949			
; INFORMATION FOR SEQ ID NO: 9:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 675 amino acids			
; TYPE: amino acid			
; TOPOLOGY: linear			
; MOLECULE TYPE: protein			
US-08-317-522A-9			
Query Match 11.8%; Score 119; DB 1; Length 675;			
Best Local Similarity 23.0%; Pred. No. 0.00025;			
Matches 47; Conservative 32; Mismatches 85; Indels 40; Gaps 8;			
Qy	11	PTVGAIPEIKIOAKAFDDIAKYFSKEWEKKASEKIFV-YMKRYEAMTKLGFATLP	69
Db	201	PGSNGEVP-KYPVTEDDVAVHFSEQWGNLSWKELYKNVMRGNYSLYSMDVAISKPD	259
Qy	70	FMCN-----KRAEFQGNLDNDPNRNQVERPQMTFGRLOGISPKIMPKPAEEG	120
Db	260	LMSQMERGERPTMQEQDSEEGETPTDPSAAH-----DGIVIKI-EVQTINDEG	306
Qy	121	NDSEEVPE--ASGPONDGKELCPGKPTTSEKIHERSGPKRGHEAWTH---RLRERKQ--	173
Db	307	SESLETPEPLMGQVEEHGFQDSSELGXPCGEQFDLDMOEPENTLEESTEGSFSELKQML	366
Qy	174	-----LVITYEISDPEDDE 188	
Db	367	VQVRNCTEGIVIKTEQEDEEBEEE 390	
RESULT 5			

```
US-08-439-818A-9
; Sequence 9, Application US/08439818A
; Patent No. 5654145
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; TITLE OF INVENTION: Trophinin and Trophinin-Assisting
; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/439,818A
; FILING DATE: 12-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/317,522
; FILING DATE: 04-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LA 1563
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 675 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-439-818A-9

Query Match 11.8%; Score 119; DB 1; Length 675;
Best Local Similarity 23.0%; Pred. No. 0.00025;
Matches 47; Conservative 32; Mismatches 85; Indels 40; Gaps 8;

QY 11 PTVGAIPEKIQKAFDDIAKYFSKEWEKMKASEKIFYV-YMKRYEAMTKLGFKATLPP 69
Db 201 PGSNGEVP-KVPVTFDDVAVHFSQEWGNLSEWQKELYKNVWNGYSELVSMYDIAISKPD 259
QY 70 FMCN-----KRAEDFQGNLDLNDNRGNQVERPQMTFGRLOGISPKIMPKKPAEEG 120
Db 260 LMSQMERGERPTMQEEDSEGETPTDPSAAH-----DGIVIKI-EVQTNDG 306
QY 121 NDSEVPE--ASGPNQDGKELCPGKPTTSEKIHSGPKRGEHAWTH---RLRERKQ-- 173
Db 307 SESLETPEPLMGQVEHGFQDSGLXPCGQEPDLDWQEPENTLESTEGSSFSFSELKQML 366
QY 174 -----LVYIEISDPDEDE 188
Db 367 VQQRNCTEGIVIKTEQEDEEEEEE 390

RESULT 6
US-08-751-965-9
; Sequence 9, Application US/08751965
; Patent No. 5858360
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; TITLE OF INVENTION: Trophinin and Trophinin-Assisting
; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

US-08-738-975-9
; Sequence 9, Application US/08738975
; Patent No. 5880267
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; TITLE OF INVENTION: Trophinin and Trophinin-Assisting
; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

US-08-751-965-9
; Sequence 9, Application US/08751965
; Patent No. 5858360
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; TITLE OF INVENTION: Trophinin and Trophinin-Assisting
; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

US-08-738-975-9
; Sequence 9, Application US/08738975
; Patent No. 5880267
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; TITLE OF INVENTION: Trophinin and Trophinin-Assisting
; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/738,975
; FILING DATE: herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/439,818
; FILING DATE: 05-Dec-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LA 2251
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 675 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-738-975-9

Query Match 11.8%; Score 119; DB 2; Length 675;
Best Local Similarity 23.0%; Pred. No. 0.00025;
Matches 47; Conservative 32; Mismatches 85; Indels 40; Gaps 8;

Qy 11 PTVGAIPEKIQKAFDDIAKYFSKEWEKMKASEKIFVY-YMKRYEAMTKLGFKATLPP 69
Db 201 PGSNGEVP-KVPVTFDDVAVHFSEQWGNLSEWQKELYKNVMRGYSLVSMYAIKPD 259
Qy 70 FMCN-----KRAEDFQGNLDNDPNRGNOVERPQMTFGRQLGISPMPKPPAE 120
Db 260 LMSQWGERPTWQEQDSEGETPTDPSAAH-----DGIWIKI-EVQTNDEG 306
Qy 121 NDSEVPE--ASGPQNDGKELCPGPKPTTSEKIHRSQPKRGEHAWTH---RLRERKQ-- 173
Db 307 SESLETPEPLMGQVEEHGFQDSSELGCPGQEPDLDMPENTLBESTGSGSEFSELKQML 366
Qy 174 -----LVYIEISDPEDDE 188
Db 367 VQQRNCTEGIVIKTEQEDEEEEEE 390

RESULT 8
US-08-728-626-9
; Sequence 9, Application US/08/728626
; Patent No. 5910451
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; TITLE OF INVENTION: Trophinin and Trophinin-Assisting
; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/728,626
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/439,818
; FILING DATE: 12-MAY-1995
```

```
;
; APPLICATION NUMBER: US 08/317,522
; FILING DATE: 04-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LA 1563
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 675 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-728-626-9

Query Match 11.8%; Score 119; DB 2; Length 675;
Best Local Similarity 23.0%; Pred. No. 0.00025;
Matches 47; Conservative 32; Mismatches 85; Indels 40; Gaps 8;

Qy 11 PTVGAIPEKIQKAFDDIAKYFSKEWEKMKASEKIFVY-YMKRYEAMTKLGFKATLPP 69
Db 201 PGSNGEVP-KVPVTFDDVAVHFSEQWGNLSEWQKELYKNVMRGYSLVSMYAIKPD 259
Qy 70 FMCN-----KRAEDFQGNLDNDPNRGNOVERPQMTFGRQLGISPMPKPPAE 120
Db 260 LMSQWGERPTWQEQDSEGETPTDPSAAH-----DGIWIKI-EVQTNDEG 306
Qy 121 NDSEVPE--ASGPQNDGKELCPGPKPTTSEKIHRSQPKRGEHAWTH---RLRERKQ-- 173
Db 307 SESLETPEPLMGQVEEHGFQDSSELGCPGQEPDLDMPENTLBESTGSGSEFSELKQML 366
Qy 174 -----LVYIEISDPEDDE 188
Db 367 VQQRNCTEGIVIKTEQEDEEEEEE 390

RESULT 9
US-08-808-599A-9
; Sequence 9, Application US/08808599A
; Patent No. 6111089
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; TITLE OF INVENTION: Trophinin, Trophinin-Assisting
; TITLE OF INVENTION: Proteins and Methods to Inhibit Implantation
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,599A
; FILING DATE: 28-FEB-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/317,522
; FILING DATE: 04-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/439,818
; FILING DATE: 12-MAY-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-LA 2256
```



```
Query Match      8.6%; Score 87.5; DB 4; Length 208;
Best Local Similarity 24.2%; Pred. No. 0.14;
Matches 47; Conservative 16; Mismatches 66; Indels 65; Gaps 8;

QY 32 FSK---EWEKMKASEKIFYVYMKRYEAMTKLGFKATLPPFMCNKRAEDFGNDLD-ND 87
Db 40 FSKKCSERWKTMSAKESKFEDMAKSDKARYDREMKNYVPP-----KGDKKKKKDPNA 93
QY 88 PNRGNQVERPQWTFRLQGISPKIMPKPA-----EKGNDSEVP----- 127
Db 94 PKR-----PPSAFFLFCSEHRPKIKSEHPGLSIGDTAKKLGEMWSEQSAKDKQPYEQKAA 148
QY 128 -----EASGPQNDGKELCPGKPTTSKIHRSKPGRGEGHAWTHRLRERKQL 174
Db 149 KLKEKYKDIAAYRAKGSKAGK--GGRPTGSKKKNEPEDEEBE----- 194
QY 175 VIYEISDPEDDE 188
Db 195 ---EEDEDEEBE 205

RESULT 13
US-09-214-881A-2
; Sequence 2, Application US/09214881A
; Patent No. 6822078
; GENERAL INFORMATION:
; APPLICANT: Ozaki, Shoichi
; APPLICANT: Sobajima, Junko
; APPLICANT: Uesugi, Hiroko
; APPLICANT: Okazaki, Takahiro
; APPLICANT: Tanaka, Masao
; APPLICANT: Nakao, Kazuwa
; APPLICANT: Yoshida, Michiteru
; APPLICANT: Shirakawa, Hitoshi
; APPLICANT: Osakada, Fumio
; TITLE OF INVENTION: DIAGNOSTIC DRUGS FOR AUTOIMMUNE DISEASES
; FILE REFERENCE: 068383.0104
; CURRENT APPLICATION NUMBER: US/09/214,881A
; CURRENT FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-214-881A-2

Query Match      8.6%; Score 87.5; DB 4; Length 208;
Best Local Similarity 24.2%; Pred. No. 0.14;
Matches 47; Conservative 16; Mismatches 66; Indels 65; Gaps 8;

QY 32 FSK---EWEKMKASEKIFYVYMKRYEAMTKLGFKATLPPFMCNKRAEDFGNDLD-ND 87
Db 40 FSKKCSERWKTMSAKESKFEDMAKSDKARYDREMKNYVPP-----KGDKKKKKDPNA 93
QY 88 PNRGNQVERPQWTFRLQGISPKIMPKPA-----EKGNDSEVP----- 127
Db 94 PKR-----PPSAFFLFCSEHRPKIKSEHPGLSIGDTAKKLGEMWSEQSAKDKQPYEQKAA 148
QY 128 -----EASGPQNDGKELCPGKPTTSKIHRSKPGRGEGHAWTHRLRERKQL 174
Db 149 KLKEKYKDIAAYRAKGSKAGK--GGRPTGSKKKNEPEDEEBE----- 194
QY 175 VIYEISDPEDDE 188
Db 195 ---EEDEDEEBE 205

RESULT 14
US-09-949-016-10728
; Sequence 10728, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10728
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10728

Query Match      8.6%; Score 87.5; DB 4; Length 320;
Best Local Similarity 24.2%; Pred. No. 0.26;
Matches 47; Conservative 16; Mismatches 66; Indels 65; Gaps 8;

QY 32 FSK---EWEKMKASEKIFYVYMKRYEAMTKLGFKATLPPFMCNKRAEDFGNDLD-ND 87
Db 152 FSKKCSERWKTMSAKESKFEDMAKSDKARYDREMKNYVPP-----KGDKKKKKDPNA 205
QY 88 PNRGNQVERPQWTFRLQGISPKIMPKPA-----EKGNDSEVP----- 127
Db 206 PKR-----PPSAFFLFCSEHRPKIKSEHPGLSIGDTAKKLGEMWSEQSAKDKQPYEQKAA 260
QY 128 -----EASGPQNDGKELCPGKPTTSKIHRSKPGRGEGHAWTHRLRERKQL 174
Db 261 KLKEKYKDIAAYRAKGSKAGK--GGRPTGSKKKNEPEDEEBE----- 306
QY 175 VIYEISDPEDDE 188
Db 307 ---EEDEDEEBE 317

RESULT 15
US-09-538-092-1193
; Sequence 1193, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuratSeqFormat Version 0.9
; SEQ ID NO 1193
; LENGTH: 654
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (0)-(0)
; OTHER INFORMATION: Polypeptide Accession Number P52736
US-09-538-092-1193

Query Match      8.5%; Score 86.5; DB 4; Length 654;
Best Local Similarity 19.5%; Pred. No. 0.93;
Matches 44; Conservative 26; Mismatches 69; Indels 87; Gaps 8;

QY 24 AFDDIAKYFSKEWEKMKASEKIFY-VYMKRYEAMTKLGFK----- 64
```

Search completed: June 20, 2005, 12:22:53  
Job time : 44 secs

**THIS PAGE LEFT BLANK**